Identification of ENF Based Grid of Origin Classification System for Media Signals: Exploring Further Features

Riyasat Ohib, Samin Yeasar Arnob, MD. Muhtady Muhaisin, Riazul Arefin, Tanzil Bin Hassan, Taslim Reza, Md. Ruhul Amin

Department of EEE, Islamic University of Technology
Board Bazar, Gazipur-1704, Bangladesh.

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 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. Recently an ENF based Machine Learning system was proposed which infers the region of recording from the power grids ENF signal extracted from the recorded multimedia signal, with the help of relevant features.

In this work, we report some features novel to this application which serve as identifying characteristics for detecting the region of origin of the multimedia recording. In addition to the ENF variation itself, the utilization of the ENF harmonics pattern extracted from the media signals as novel features enables a more accurate identification of the region of origin. These characteristics were used in a multiclass machine learning implementation to identify the grid of the recorded signal.

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

1. Introduction

The supply frequency of the power distribution network i.e. ENF has a nominal value of 50/60 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 use of ENF signal to identify modified audio recordings was proposed by Grigoras [2-5] presently being used in multimedia forensics applications, as it gets embedded in the multimedia recordings made in the vicinity of electrical activity. The ENF signals can also be used to identify and classify video signals as well [6]. 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 [7]. 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.

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 [8].

Before the advent of Machine Learning techniques, the sole focus was using the unique ENF pattern of the power grids as sole criteria to differentiate between the media recording signals originating from different power grids. As different grids generated unique ENF signals those signals were extracted from the power signals and chronologically stored for future reference [3, 9]. Whenever an audio file needed to be classified they were cross-referenced by correlating them with the database [3]. The Machine Learning approach greatly alleviates these problems to some extent. Moreover, we found that a Machine Learning approach also lets us exploit other differentiating features between different Grids which used in tandem with the features extracted from ENF signals provide an enhancement in their ability to classify. We found that feature-sets for the support vector machine can be developed that does not arise from the ENF pattern itself, but from the temporal media signal. This paper proposes the utilization of the 'Fourier Transform Profile' or the pattern of the harmonics of the ENF signals in the frequency domain as novel features.

Our database [10] contains power and audio recordings from 12 different grids around the world, which includes both noise-free power recordings and noisy audio recordings sampled at 1000Hz. We have also extracted Power recordings with the help of sensing circuit from National Power Grid of Bangladesh and used it alongside the other grids available. Using statistical features derived from the ENF signals, already discussed in [8], as well as proposed novel features mentioned in this paper, the region of recording of the media signals were differentiated based on a Machine Learning approach.

2. ENF Extraction and Database Description

Media recordings of 12 different grid from all over the world were available on an online database [10] published by the authors of [8] which is subset of the data used in their work. As the whole dataset was not published by the authors, an exact comparative analysis with their work could not be performed in this paper. However the ENF extraction code used by the authors of [8] was available, therefore a comparative study using their extraction algorithm and code was performed on their features compared to our proposed features in the publicly available dataset. This demonstrates the effects of the novel features on the classification efficiency in a standard extraction algorithm and classification system. As the extraction code are exactly the same, the only difference between the two works are in the additional features used in this work and in the coding platform of classification system used - the previous work using LIBSVM and this work using the MATLAB ECOC multiclass classification system.

Each of the grid contained Power and Audio data. We have also extracted Power recordings of National Power Grid of Bangladesh for analysis. The number of available examples for Power and Audio recordings of each grid, nominal frequencies corresponding to each grid and their range of variation are provided in Table-1. The audio files from Grids J to M were not at our disposal. The

Database contained variable numbers of power recordings but same number of audio recordings across all grids. Main recording of each power and audio file was at first divided into segments of ten minutes to create substantial number of datasets so that we can train our program properly for future prediction. Each of the segments were divided into non-overlapping frames of 5 seconds duration & then Spectrum Combining method [11] was used for estimating the dominant frequency component of each time frame, thus generating the ENF signal segment of length S = 120 samples.

Table 1. Description of the database

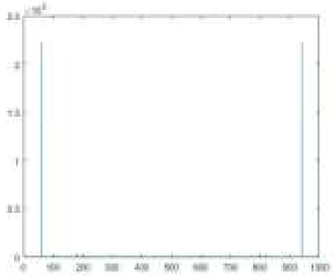
Grid	No. of Power Examples	No. of Audio Examples	Nominal Frequency (Hz)	Maximum Frequency (Hz)	Minimum Frequency (Hz)	Frequency Range (Hz)
Texas (A)	33	6	60	60.032	59.651	0.067
Lebanon (B)	38	6	50	50.580	49.142	1.707
Eastern U.S (C)	34	6	60	60.044	59.962	0.082
Turkey (D)	37	6	50	50.057	49.942	0.115
Ireland (E)	38	6	50	50.077	49.925	0.152
France (F)	26	6	50	50.778	49.942	0.134
Tenerife (G)	36	6	50	50.142	49.801	0.341
India (Agra) (H)	41	6	50	50.316	49.716	0.600
Western U.S (I)	37	6	60	60.047	59.951	0.095
Brazil (J)	35	-	60	60.0655	59.8662	0.1993
Norway (K)	35	-	50	50.1069	49.8734	0.2335
Australia (L)	36	-	50	50.0963	49.8701	0.2262
Bangladesh (M)	49	-	50	51.2637	49.9932	2.2705

3. Feature Analysis

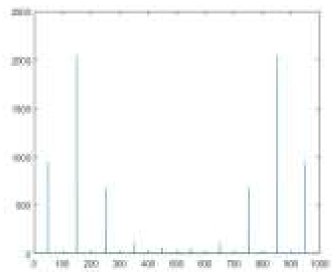
From the empirical differences in the variations of the ENF signals, we can extract meaningful statistical features for our classification system. We took a set of ENF signal segments of fixed size S = 120 samples from among the power Grids of our dataset. We adopted the 16 features included in [8] and introduced 16 more features, all of which arises from the Fourier Transform profiles of the audio and power examples.

An auto regressive (AR) statistical model was used similar to the one in [8]. However, we explored the possibility of using higher orders of AR models to improve the performance of the classification system. To select the order of the AR model, the approximations to Schwarz's Bayesian Criterion and to the logarithm of Akaike's Final Prediction Error were computed [12, 13]. These computations revealed that a second order AR model is indeed the most optimized model for this dataset, thus corroborating the choice of AR model in [8].

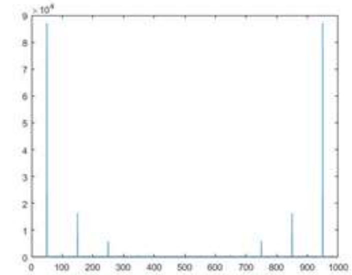
3.1 Fourier Transform Profile



1(a)



1(b)



1(c)

Fig 1: Magnitude of the harmonic components of power recording from (a) Grid A, (b) Grid D, (c) Grid G.

It was empirically evident that the magnitude of the harmonic components of the power and audio recordings uniquely differs from one grid to another in general and the Fourier Transformed frequency domain data had significant variations in their magnitude at different frequencies in the recordings from one grid to another and showcased common trends Fig.1.

Whenever a non-linear device or load is subjected to a source of sinusoidal voltage, the resulting current is not perfectly sinusoidal. Consequently this gives rise to harmonics at the load terminals [14].

Even though there are standards to deal with power quality issues such as harmonics and inter-harmonics [15] and for specific equipment dealing under the influence of harmonics [16-18], individual countries make their own adjustments to accommodate for their unique power consumption scenarios and national priorities, motivated by special characteristics of their power system configuration and load management systems (e.g. the use of ripple control in some countries) [14]. The random difference of load and various harmonics control equipment at different grids and nations most likely give rise to underlying statistical patterns in the harmonics of the clean power and audio recordings in a manner similar to that of the statistical patterns of the ENF. Further research is needed to fully identify the origins of the distinctions in the Fourier Transform profile and their similarity in the recordings from the same grid and to identify whether recording equipment play any role whatsoever and to what degree.

To exploit this underlying property, the normalised magnitudes of the harmonic components of the ENF signal from the audio and power data were extracted. These are directly obtained by applying the Fourier transformation on the time-domain power and audio signals and extracting the magnitudes of the ENF harmonics as an eight element vector. A polynomial model of the 7^{th} degree was fit over the normalized magnitudes of the harmonic components thus extracted and the coefficients and the constant term of the fitted polynomial was used as distinguishing features termed as ‘Polynomial Model Parameters’ (features 17-24). Moreover, the ‘Normalized magnitudes (amplitudes)’ of the harmonic components itself was also utilized as features (25-32). The order of polynomial was set on the basis of the number of points that the polynomial fits on. There are 8 harmonic components which result in 8 normalized magnitude points, thus a 7-order polynomial fit perfectly traces all the points.

The features are added sequentially & their impact on classification efficiency are summarized in Table-2.

Table 2. Used Feature Sets

Feature Number	Feature sets	
1	Mean of ENF segment	Feature Used in the Work [8]
2	Log (variance) of ENF segment	
3	Log(range) of ENF segment	
4	Log(variance) of approximation after F-level wavelet analysis (F=9)	
5-13	Log(variance) of nine levels of detail signals computed through F-level wavelet analysis from coarser to finer (F=9)	
14-15	AR(2) model parameters	
16	Log(variance) of the innovation signal after AR(2) modelling	
17-24	Polynomial model parameter	Proposed Addition Features
25-34	Normalized Amplitude of the Harmonic components	

3.2 Trained Classification Systems

While training the Classification System the normalization parameters were stored, and later used to normalize the Testing Data. We resorted to a similar technique as the one reported in [8]. In this work, a multiclass classifier based on the Error-correcting output codes (ECOC) multiclass model using support vector machine (SVM) binary learners on the MATLAB platform [19-22]. This employs a one vs one model. A Gaussian Radial Basis kernel function with automatic kernel scaling was used. When the Kernel scale mode is set to 'Auto', MATLAB uses a heuristic procedure to select the scale value; the heuristic procedure uses subsampling [23].

4. Results

In this section, we present the results obtained from our analysis. Each of the proposed features affects the ability of the Learning System to detect a region of origin of the multimedia file.

For calculation of efficiency of classification, Machine Learning and Statistics Toolbox functions was used. 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. MATLAB then trains 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.

For the classification system trained only on power recording we achieved an accuracy of 96.03% with a 20-fold Cross-Validation scheme selecting all the features. Change in efficiency with feature inclusion is shown in Table 3. Classification accuracy of power files rose from 77.7% to 89.83% after the inclusion of features (17-24) and further addition of the maximum value of the normalized Amplitude of Harmonic Components (25-32) boosts the efficiency to 96.03%.

For the classification system trained only on power recording we achieved an accuracy of 88.90% with a 20-fold Cross-Validation scheme selecting all the features. The classification accuracy of audio files first changes from 74.70% to 77.00% with the inclusion of features (17-24) and finally rises to 88.90% after addition of features (25-32).

Table 3. Change in Efficiency for Feature Sets Inclusion

Feature no.	Feature inclusions	Efficiency (Audio)	Efficiency (Power)
1-16	Original 16 Features	74.70%	77.7%
17-24	Polynomial model parameter	77.00%	88.47%
25-32	Normalized Amplitude of the Harmonic components	88.90%	96.03%

To closely understand the results of our approach the confusion matrix is presented in Table 4. Here, the efficiency of our system in predicting various grids are presented. For system trained on Power files, the (E) Ireland has the lowest efficiency of identification of 88%. The highest efficiency of 100% is achieved in identification of Grids of (B) Lebanon, (C) Eastern US, (H) India, (J) Brazil, (K) Norway, (M) Bangladesh. Apart from the Ireland grid, for all the other grids, the system has an identification accuracy above of 90%. The relatively low identifying accuracy for the Ireland is understandable as the system confuses some recordings from the Ireland grid which is 50Hz group to some other grids with 50Hz frequency. This is easily observable from the confusion matrix (Table 4). In contrast, the highest identifying efficiency was achieved mostly for grids whose ENF signals are significantly different from one another.

Table 4. Confusion Matrix of Power Examples

Testing Classes	No. of Examples	Grid A	Grid B	Grid C	Grid D	Grid E	Grid F	Grid G	Grid H	Grid I	Grid J	Grid K	Grid L	Grid M
(A) Texas	33	92%	-	-	-	-	-	-	-	-	-	-	-	-
(B) Lebanon	38	-	100%	-	-	5%	-	-	-	-	-	-	-	-
(C) Eastern U.S	34	8%	-	100%	-	-	-	-	-	5%	-	-	-	-
(D) Turkey	37	-	-	-	95%	2%	6%	2%	-	-	-	-	-	-
(E) Ireland	38	-	-	-	3%	88%	3%	-	-	-	-	-	-	-

(F) France	26	-	-	-	-	2%	91%	-	-	-	-	-	-	-
(G) Tenerife	36				3%	2%		98%	-	-	-	-	-	-
(H) India (Agra)	41	-	-	-	-	-	-	-	100%	-	-	-	-	-
(I) Western U.S	37	-	10%	-	-	-	-	-	-	95%	-	-	-	-
(J) Brazil	35	-	-	-	-	-	-	-	-	-	100%	-	-	-
(K) Norway	35	-	-	-	-	-	-	-	-	-		100%	3%	-
(L) Australia	36	-	-	-	-	-	-	-	-	-	-	-	97%	-
(M) Bangladesh	58	-	-	-	-	-	-	-	-	-	-	-	-	100%

For system trained on Audio files, the highest efficiency of 100% is achieved in identification of almost all of the grids except Grids of (C) Eastern US (83%), (H) India (83%), (E) Ireland (60%). We have seen from both Audio & Power file that our feature sets make bad prediction in identifying Grid of Ireland and that severely affects in overall efficiency. This is easily observable from the confusion matrix (Table 5).

Table 5. Confusion Matrix of Audio Examples

Testing Classes	No. of Examples	Grid A	Grid B	Grid C	Grid D	Grid E	Grid F	Grid G	Grid H	Grid I
(A) Texas	6	100%	-	-	-	-	-	-	17%	-
(B) Lebanon	6	-	100%	-	-	-	-	-	-	-
(C) Eastern U.S	6	-	-	83%	-	10%	-	-	-	-
(D) Turkey	6	-	-	-	100%	10%	-	-	-	-
(E) Ireland	6	-	-	-	-	60%	-	-	-	-
(F) France	6	-	-	17%	-	-	100%	-	-	-
(G) Tenerife	6	-	-	-	-	-	-	100%	-	-
(H) India (Agra)	6	-	-	-	-	10%	-	-	83%	-
(I) Western U.S	6	-	-	-	-	10%	-	-	-	100%

In Table 3, the comparison between the identification efficiencies using only the 16 features presented in [8] and with addition of our proposed features, are presented. Even though the dataset used in this work is a subset of the one used in [8], as the same extraction code published by the authors [24] was used in both the works the only difference in identification efficiency is due to the feature set used and the platform used. Both the LIBSVM, used in the previous work and MATLAB's ECOC multiclass system uses a one vs one multiclass model, weighted SVM, and works in a similar way [20, 21, 25].

The lack of data in comparison to the work in [8] explains the somewhat reduced accuracies of this program while using the original feature-set of 16 features, as more training data generally leads to better approximations and accuracy of the machine learning system [26].

For system trained on power recordings including all features of this work, the average classification efficiency over 20 rounds of trials 96.03% was achieved. In a system trained on only audio file, average classification efficiency of 88.90% was achieved, which are a significant improvement over the efficiency achieved by the original 16 feature set on the same algorithm, same datasets and classification platform.

Table 6. Identification Efficiency

Trained System	Classification Efficiency (Power)	Classification Efficiency (Audio)
Power	96.03%	-
Audio	-	88.90%

5. Conclusion

In this work, we implemented a Machine Learning System to identify the Grids of origin of an ENF signal and explored some novel classification features that might aid the Machine Learning systems ability to identify the grids of origin. 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. This work also demonstrates the applicability of MATLABs ECOC Multiclass Classification System in classifying the region of origin of multimedia recordings through the analysis of ENF signals embedded in them. Finally, we have managed to increase the identification efficiency of power and audio recording through this work to 96.03 % and 88.90% respectively, with the Fourier Transform Profile feature acting as a significant catalyst for boosting up the identifying efficiency.

References

1. Bollen, M.H. and I. Gu, *Signal processing of power quality disturbances*. Vol. 30. 2006: John Wiley & Sons.
2. Grigoras, C., *Digital audio recording analysis–the electric network frequency criterion.* International Journal of Speech Language and the Law, 2005. **12**(1): p. 63-76.
3. Grigoras, C., *Applications of ENF criterion in forensic audio, video, computer and telecommunication analysis.* Forensic Science International, 2007. **167**(2): p. 136-145.
4. Grigoras, C. *Applications of ENF Analysis Method in Forensic Authentication of Digital Audio and Video Recordings*. in *Audio Engineering Society Convention 123*. 2007: Audio Engineering Society.
5. Grigoras, C., *Applications of ENF analysis in forensic authentication of digital audio and video recordings.* Journal of the Audio Engineering Society, 2009. **57**(9): p. 643-661.
6. Garg, R., A.L. Varna, and M. Wu. *Seeing ENF: natural time stamp for digital video via optical sensing and signal processing*. in *Proceedings of the 19th ACM international conference on Multimedia*. 2011: ACM.
7. Maher, R.C., *Audio forensic examination.* Signal Processing Magazine, IEEE, 2009. **26**(2): p. 84-94.

8. Hajj-Ahmad, A., R. Garg, and M. Wu, *ENF-Based Region-of-Recording Identification for Media Signals.* Information Forensics and Security, IEEE Transactions on, 2015. **10**(6): p. 1125-1136.
9. Cooper, A.J. *The electric network frequency (ENF) as an aid to authenticating forensic digital audio recordings–an automated approach*. in *Audio Engineering Society Conference: 33rd International Conference: Audio Forensics-Theory and Practice*. 2008: Audio Engineering Society.
10. Adi Hajj-Ahmad, M.W. *IEEE Signal Processing Cup 2016, DATASETS* 2016; Available from: http://spcup16.ece.umd.edu/datasets.php.
11. Hajj-Ahmad, A., R. Garg, and M. Wu, *Spectrum combining for ENF signal estimation.* Signal Processing Letters, IEEE, 2013. **20**(9): p. 885-888.
12. Neumaier, A. and T. Schneider, *Estimation of parameters and eigenmodes of multivariate autoregressive models.* ACM Transactions on Mathematical Software (TOMS), 2001. **27**(1): p. 27-57.
13. Schneider, T. and A. Neumaier, *Algorithm 808: ARfit—A Matlab package for the estimation of parameters and eigenmodes of multivariate autoregressive models.* ACM Transactions on Mathematical Software (TOMS), 2001. **27**(1): p. 58-65.
14. Arrillaga, J. and N.R. Watson, *Power system harmonics*. 2004: John Wiley & Sons.
15. 61000-2-1, I., *Electromagnetic Compatibility (EMC)—Part 2: Environment—Section 1: Description of the Environment–Electromagnetic Environment for Low-Frequency Conducted Disturbances and Signalling in Public Power Supply Systems.* IEC, Geneva., 1990.
16. Telegraph, I.T.U.I. and T.C. Committee, *Directives Concerning the Protection of Telecommunication Lines Against Harmful Effects from Electricity Lines: Text Approved by the 2nd Plenary Assembly of the International Telegraph and Telephone Consultative Committee CCITT, New Dehli 1960*. 1963: International Telecommunication Union.
17. C57.110-1998, I.A., *IEEE recommendation practice for establishing transformer capability when supplying nonsinusoidal load currents.* IEEE, New York.
18. IEEE/ANSI Std 18–2002 (Reaff 1991), *IEEE Standard for Shunt Power Capacitors.* IEEE, New York.

.

19. Allwein, E.L., R.E. Schapire, and Y. Singer, *Reducing multiclass to binary: A unifying approach for margin classifiers.* Journal of machine learning research, 2000. **1**(Dec): p. 113-141.
20. Escalera, S., O. Pujol, and P. Radeva, *Separability of ternary codes for sparse designs of error-correcting output codes.* Pattern Recognition Letters, 2009. **30**(3): p. 285-297.
21. Escalera, S., O. Pujol, and P. Radeva, *On the decoding process in ternary error-correcting output codes.* IEEE transactions on pattern analysis and machine intelligence, 2010. **32**(1): p. 120-134.
22. Fürnkranz, J., *Round robin classification.* The Journal of Machine Learning Research, 2002. **2**: p. 721-747.
23. MathWorks, M. *Choose Classifier Options*. 2016 [cited 2017 31 March]; Available from: https://www.mathworks.com/help/stats/choose-a-classifier.html#bunt0n0-1.
24. Adi Hajj-Ahmad, R., MinWu,. *Spectrum Combining for ENF Signal Estimation - Dataset & Codes,*. 2013 [cited 2017 March 03]; Available from: http://cwwongpatch2.no-ip.org/mast_new/research.php?t=enf.

25. Chang, C.-C. and C.-J. Lin, *LIBSVM: a library for support vector machines.* ACM Transactions on Intelligent Systems and Technology (TIST), 2011. **2**(3): p. 27.
26. Banko, M. and E. Brill. *Scaling to very very large corpora for natural language disambiguation*. in *Proceedings of the 39th annual meeting on association for computational linguistics*. 2001: Association for Computational Linguistics.