ON ROBUST ELECTRIC NETWORK FREQUENCY DETECTION USING HUBER REGRESSION

Christos Korgialas
Aristotle University of Thessaloniki
Thessaloniki, Greece
ckorgial@csd.auth.gr

Constantine Kotropoulos
Aristotle University of Thessaloniki
Thessaloniki, Greece
costas@csd.auth.gr

ABSTRACT

A robust regression technique known as Huber regression is incorporated into the Electric Network Frequency (ENF) detection task. This novel framework is based on the assumption of a mixture noise model, which combines Gaussian and Laplacian noise for ENF detection in short-length audio recordings. The effectiveness of the proposed ENF detector is assessed through accuracy calculations and the analysis of the Receiver Operating Characteristic curve with respect to the Area Under the Curve. Real-world benchmark data from the ENF-WHU dataset are utilized for this evaluation. The experimental results indicate that integrating the Huber regression method leads to a significant enhancement in ENF detection for short-length audio recordings, outperforming the performance of existing state-of-the-art techniques.

CCS CONCEPTS

• **Mathematics of computing** → **Robust regression**; • **Applied computing** → *Investigation techniques.*

KEYWORDS

Electric Network Frequency (ENF), Huber Regression, Likelihood Ratio Detector (LRT), Huber-LRT Detector, Robust ENF Detection, Multimedia Forensics

ACM Reference Format:
Christos Korgialas and Constantine Kotropoulos. 2023. ON ROBUST ELECTRIC NETWORK FREQUENCY DETECTION USING HUBER REGRESSION. In *27th Pan-Hellenic Conference on Progress in Computing and Informatics (PCI 2023), November 24–26, 2023, Lamia, Greece.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3635059.3635098

1 INTRODUCTION

The Electric Network Frequency (ENF), which originates from the fluctuations of the power grid frequency, serves as a distinctive and intrinsic "fingerprint" within multimedia content, such as audio recordings [2]. ENF holds a nominal value of 50 Hz in Europe and 60 Hz in the United States/Canada, serving as a forensic criterion in the realm of multimedia forensics [9, 13, 25]. The comparison of the estimated ENF from a multimedia recording with a ground truth ENF derived from power mains can find application in tasks such as time stamp verification [5, 6, 12, 14, 35], geo-localization [4, 32, 33], and tampering detection [11, 28, 31, 34]. Additionally, the integration of ENF analysis with Non-Intrusive Load Monitoring techniques [3, 21, 30] further expands forensic capabilities, allowing for the disaggregation of energy consumption patterns that may corroborate the time and date of a recording.

To achieve optimal performance in the aforementioned applications, an important prerequisite before ENF estimation is the detection of ENF signals in multimedia recordings. A significant contribution that tackles ENF detection is developed in [15]. Six distinct detectors are developed and assessed, with three of them specifically tested using real-world data alongside synthetic data. The detectors are evaluated, encompassing both long-length audio recordings and short-length audio recordings. For ENF detection in short-length audio recordings, a detector is proposed to enhance the Likelihood Ratio Test (LRT) performance employing the Least Absolute Deviations (LAD) regression [23]. LAD regression operates under the assumption of a Laplacian noise model, solving for regression weights with fixed frequency estimates and solving for frequency estimates with constant regression weights until achieving convergence. In [24], a multi-tone time-frequency detector is developed employing a combination of multiple harmonics to identify the presence of valid ENF traces within a recording. Additionally, this detector provides insights into the overall quality of the ENF signal and the count of available harmonic components. An ENF detector employing a superpixel approach for videos is developed in [29]. This method involves ENF signal estimations derived from stable superpixel regions to determine the presence or absence of an ENF signal in brief video clips. In [27], a linear discriminant is employed to create an automated detector for ENF disturbances. Prior to assessing the detector, ENF extraction is carried out using the Estimation of Signal Parameters by Rotational Invariant Techniques.

Significant research is also directed toward the ENF estimation task. A multi-tone harmonic model for estimating the ENF is introduced in [1]. To enhance the accuracy of ENF signal estimation, multiple harmonics are combined, and the Cramer-Rao bound is utilized to restrict the variance of the ENF estimator. In [10], a spectral estimation method is presented that integrates the ENF across various harmonics. The extraction of ENF takes into account the local signal-to-noise ratio at each harmonic. Instead of using the conventional Short-Time Fourier Transform (STFT) [26], the process of ENF extraction is approached as a data-dependent filtering issue. A framework for robust extraction of ENF from real audio recordings, which encompasses multi-tone ENF harmonic enhancement and the utilization of a graph-based approach to optimize harmonic selection, is introduced in [16]. Drawing insights from both [7] and [8], the frequency demodulation process in [18] utilizes the intermediate frequency signal's spectrum to deduce the

PCI 2023, November 24–26, 2023, Lamia, Greece

ACM ISBN 979-8-4007-1626-3/23/11.
https://doi.org/10.1145/3635059.3635098

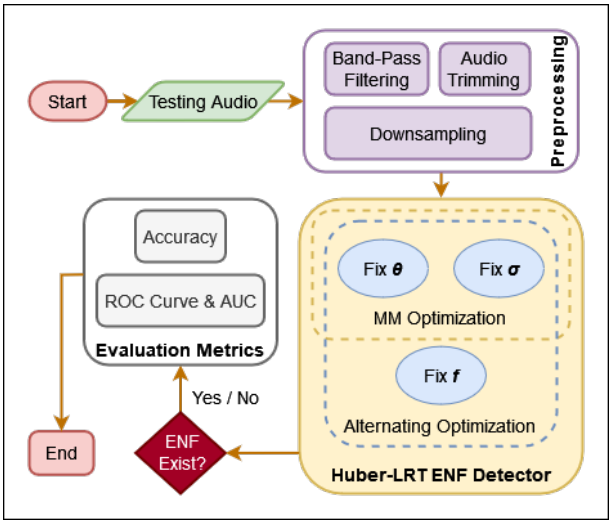


Figure 1: The proposed Huber-LRT ENF detection architecture.

highest achievable frequency of the ENF. In [19], filter-bank Capon spectral estimators and non-rectangular temporal windowing are developed to enhance ENF estimation accuracy for authenticity verification. Moreover, a non-parametric approach that embeds a customized lag window into the Blackman-Tukey method is proposed in [20], reducing speech content interference and improving forensic analysis precision. The evaluation of ENF extraction involves the assessment of non-parametric and parametric spectral estimation techniques.

Here, ENF detection is approached through the perspective of Huber regression [17]. A robust regression technique is proposed that combines the strength of the ℓ_2 norm in Least Squares (LS) and the robustness of the ℓ_1 norm in LAD regression. The proposed method incorporates a novel strategy by assuming a combination of Gaussian and Laplacian noise models within a mixture noise model framework. This is insightful, as the distribution of the noise model is similar to a Gaussian distribution in its central tendency while exhibiting a double exponential distribution in its tails. By employing a mixture noise model, a balance is obtained between the smoothness of Gaussian noise and the heavy-tailed behavior of Laplacian noise. This approach advances ENF detection, surpassing the robust ENF detector in [23] and enabling enhanced detection accuracy. By integrating Huber regression with a Gaussian-Laplacian mixture noise model, the proposed method leads a revolutionary stride in ENF detection, achieving exceptional accuracy and versatility compared to state-of-the-art techniques.

The rest of the paper is structured as follows. In Section 2, the proposed framework is analyzed, while in Section 3, experimental findings are presented. Section 4 concludes the research approach while also providing directions for future research.

2 PROPOSED FRAMEWORK

In this Section, the proposed framework, depicted in Figure 1, is analyzed. A description of the signal model as the formulation of the problem is described in Section 2.1, while the proposed ENF detection framework is discussed in Section 2.2.

2.1 Signal Model

The ENF is considered a deterministic signal denoted as $s[n]$ in the presence of noisy observations $x[n]$. The ENF waveform $s[n]$ can be expressed as:

$$s[n] = A[n]\ \cos\left(2\pi\ T\, f[n] + \phi\right),\ n \in \{0, 1, \ldots, N-1\}, \quad (1)$$

where $A[n] > 0$ and $f[n]$ represent the time-varying amplitude and ENF frequency respectively. The parameter ϕ signifies the unknown initial phase. Here, $T = 1/f_s$ corresponds to the sampling interval, with f_s denoting the sampling frequency.

Due to the slow variation of the ENF signal over time, $A[n]$ and $f[n]$ can be approximated by constants A_c and $f_c\, n$, respectively, resulting in the simplified expression of the ENF waveform

$$s_c[n] = A_c\ \cos(2\pi\, T\, f_c\, n + \phi). \quad (2)$$

The task of detecting ENF involves a binary hypothesis scenario. This binary framework facilitates the precise detection of the ENF presence within the signal model as follows:

$$\begin{aligned} \mathcal{H}_0 &: x[n] = w[n] \\ \mathcal{H}_1 &: x[n] = s_c[n] + w[n], \end{aligned} \quad (3)$$

where $w[n]$ represents the assumed combined noise model resulting from the aggregation of Gaussian noise $u_{\text{Gaussian}}[n]$ and Laplacian noise $u_{\text{Laplacian}}[n]$. Moreover, $x[n]$ corresponds to preprocessed samples of the time-domain signal. The preprocessing stage (see Section 3.2) consists of a threefold process, each serving a specific purpose. Bandpass filtering is applied to retain the ENF signal while removing unwanted noise. Downsampling is performed to reduce the sample rate, making the data more manageable and efficient for analysis. Finally, audio trimming is employed to preserve the short-length audio recordings, each lasting 5 seconds in duration.

2.2 Huber-LRT ENF Detection

The core of the detection problem centers on the assumption of a mixed noise model. This is where the integration of Huber regression for ENF detection stands with its alignment to the mixed noise model's complexity. By incorporating both ℓ_1 norm and ℓ_2 norm elements, Huber regression adapts to the varied noise patterns inherent in the mixed noise model. The ℓ_1 norm component equips the method to resiliently manage Laplacian noise, which handles outliers robustly [36]. Simultaneously, the inclusion of the ℓ_2 norm element empowers Huber regression to handle the Gaussian noise aspects within the mixture, aligning with the principles of least squares.

ENF detection initiates with estimating the maximum likelihood values [22] for the unknown parameters, achieved through solving the optimization problem depicted as:

$$\left\{\hat{f}_c,\, \hat{\boldsymbol{\theta}},\, \hat{\sigma}\right\} = \underset{f_c, \boldsymbol{\theta}, \sigma}{\operatorname{argmin}}\, J_{\text{Huber}}\left(f_c,\, \boldsymbol{\theta},\, \sigma\right), \quad (4)$$

where $\boldsymbol{\theta} \in \mathbb{R}^{2\times 1}$ denotes the regression coefficients and $\sigma > 0$ the regression scale. The objective function, grounded in Huber regression, is articulated as:

$$J_{\text{Huber}}(f_c, \boldsymbol{\theta}, \sigma) = \frac{2N}{2}(\mu\sigma) + \rho\left(\frac{\boldsymbol{x} - \boldsymbol{H}(f_c)\,\boldsymbol{\theta}}{\sigma}\right)\sigma. \tag{5}$$

In (5), $\boldsymbol{x} = [x[0], x[1], \ldots, x[N-1]]^\top$, N is the number of samples in $\boldsymbol{x}$, and $[\cdot]^\top$ denotes the transpose operator. Additionally, $\rho : \mathbb{R}^N \to \mathbb{R}$ is a convex and differentiable loss function, which embodies a duality:

$$\rho(\boldsymbol{k}) = \frac{1}{2}\begin{cases} \|\boldsymbol{k}\|_2, & \text{for } \|\boldsymbol{k}\|_2^2 \le c^2 \\ 2c\|\boldsymbol{k}\|_1 - c^2, & \text{for } \|\boldsymbol{k}\|_1 > c, \end{cases} \tag{6}$$

where $\boldsymbol{k}$ within ρ corresponds to the term $\frac{\boldsymbol{x}-\boldsymbol{H}(f_c)\,\boldsymbol{\theta}}{\sigma}$. The Huber loss function is a hybrid error measure that combines the quadratic penalty of the ℓ_2 norm (i.e., LS loss) for small errors with the linear penalty of the ℓ_1 norm (i.e., LAD loss) for large errors, providing a robust and balanced approach to error minimization. Moreover, $\mu > 0$ ensures Fisher-consistency of σ under i.i.d. Gaussian errors, calculated through:

$$\mu = \frac{1}{2}\, c^2 \left(1 - F_{\chi_1^2}\left(c^2\right)\right) + F_{\chi_3^2}\left(c^2\right), \tag{7}$$

where $F_{\chi_k^2}$ is the cumulative χ_k^2 distribution, while $c = 1.345$ stands for a user-defined tuning threshold influencing the robustness level. In (5), the matrix $\boldsymbol{H}(f_c) = [\boldsymbol{\alpha}|\boldsymbol{\beta}] \in \mathbb{R}^{N\times 1}$ is composed of columns:

$$\begin{aligned} \boldsymbol{\alpha} &= \left(1, \cos\left(2\pi\, T, f_c \cdot 1\right), \ldots, \cos\left(2\pi\, T, f_c \cdot (N-1)\right)\right)^\top \\ \boldsymbol{\beta} &= \left(0, \sin\left(2\pi\, T, f_c \cdot 1\right), \ldots, \sin\left(2\pi\, T, f_c \cdot (N-1)\right)\right)^\top. \end{aligned} \tag{8}$$

The optimization problem (4) involves iterative estimation of f_c, regression coefficients $\boldsymbol{\theta}$, and the scale parameter σ. The iteration begins by setting f_c to the frequency corresponding to the peak of the periodogram, computed as:

$$\hat{f_c} = \underset{f}{\text{argmax}} \left| \sum_{n=0}^{N-1} x[n] e^{-j2\pi T f n} \right|^2. \tag{9}$$

Then the parameters $\boldsymbol{\theta}$ and σ are estimated by solving the Huber regression problem:

$$\left\{\hat{\boldsymbol{\theta}}, \hat{\sigma}\right\} = \underset{\boldsymbol{\theta},\sigma}{\text{argmin}} \left\{ \frac{2N}{2}(\mu\sigma) + \rho\left(\frac{\boldsymbol{x} - \boldsymbol{H}(\hat{f_c})\,\boldsymbol{\theta}}{\sigma}\right)\sigma \right\}, \tag{10}$$

where (10) is convex when considering both the regression vector and the scale parameter jointly. This convexity property is based on the assumption that the loss function $\rho(\cdot)$ used is itself convex. The estimated parameters $\hat{\boldsymbol{\theta}}$ and $\hat{\sigma}$ result through the utilization of a block-wise Minimization-Majorization (MM) algorithm [36]. The MM algorithm iteratively seeks to minimize the Huber loss (6) by constructing a sequence of surrogate functions that majorize the original Huber loss function. These surrogate functions are typically chosen to be quadratic approximations near the current parameter estimates, making the optimization more tractable. At each iteration, the MM algorithm updates the parameter estimates by minimizing the surrogate function, with the goal of minimizing the original Huber loss. This iterative process continues until a convergence criterion is met, resulting in parameter estimates that minimize the Huber loss criterion (10). Subsequently, after obtaining the parameter estimates $\hat{\boldsymbol{\theta}}$ and *sigma*, these values are held fixed, and the estimation of the frequency $\hat{f}$ is pursued through a separate optimization problem:

$$\hat{f} = \underset{f}{\text{argmin}} \left\{ \frac{2N}{2}(\mu\hat{\sigma}) + \rho\left(\frac{\boldsymbol{x} - \boldsymbol{H}(f)\,\hat{\boldsymbol{\theta}}}{\hat{\sigma}}\right)\hat{\sigma} \right\}. \tag{11}$$

The minimization of (11) is achieved by implementing a dense grid search methodology around the ENF center frequency $\hat{f_c}$. Following the completion of the grid search and upon obtaining the initial solution for the frequency f, the optimization problem specified in (10) is subjected to further refinement. This refinement is achieved through an iterative alternating optimization approach. The primary objective of this iterative process is to attain convergence of the frequency parameter to its optimal value $\hat{f}$, while concurrently optimizing the parameters $\boldsymbol{\theta}$ and σ using the Huber regression method. While the optimization problem in (10) is inherently convex and can be efficiently solved, the grid search in (11) may lead to suboptimal results.

Once the unknown parameters $\hat{f}$, $\hat{\boldsymbol{\theta}}$, and $\hat{\sigma}$ have been estimated, a Huber-LRT detector is established to assess whether $\boldsymbol{x}$ falls into the $\mathcal{H}_1$ scenario relative to a threshold denoted as η, i.e.,

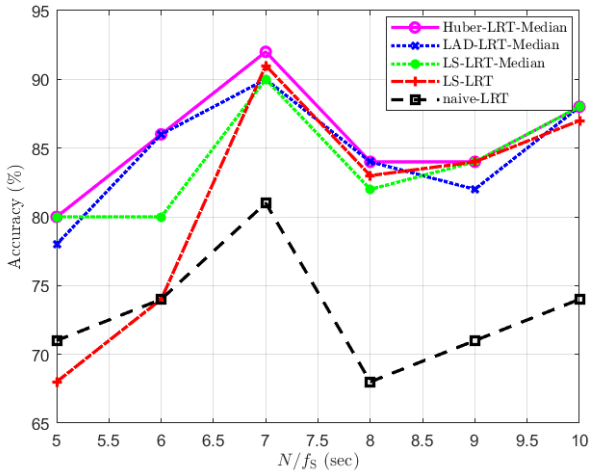
$$T_{\text{Huber}}(\boldsymbol{x}) = \frac{\boldsymbol{x}^\top \boldsymbol{H}(\hat{f})\,\hat{\boldsymbol{\theta}}}{\boldsymbol{x}^\top \boldsymbol{x}} > \eta. \tag{12}$$

The threshold for the Huber-LRT detector is determined, following the process in [23], by computing the median of the test statistic values for each duration across all recordings under both $\mathcal{H}_0$ and $\mathcal{H}_1$. Let M be the number of recordings for each duration, which in this case is $M = 100$. This is in accordance with the 100 audio recordings available for analysis, which are divided into two groups, each containing 50 recordings (see Section 3.1). When M is even, the threshold is calculated as follows:

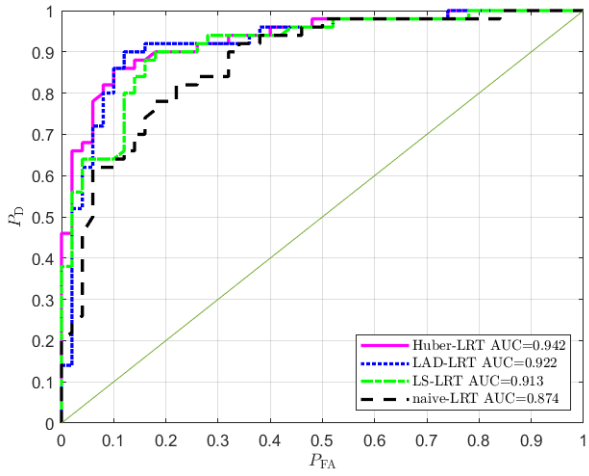
$$\eta = \frac{1}{2}\left(T_{\text{Huber},\,(\frac{M}{2})} + T_{\text{Huber},\,(\frac{M}{2}+1)}\right), \tag{13}$$

Table 1: Summary of preprocessing steps and values for ENF detection.

Preprocessing Parameter	Value
Duration Range	5 to 10 seconds
Downsampling	44.1 to 8 kHz
Resampling	400 Hz
Bandpass Filtering	
- Center Frequency	100 Hz
- Cut-off Frequencies	[99.9, 100.1] Hz
- Transition Bandwidth	0.1 Hz
- Pass-Band Alignment	2^{nd} ENF harmonic (100 Hz)



(a) Detection accuracy across different recording durations.



(b) ROC curves and corresponding AUC values for 5-second duration recordings.

Figure 2: Performance comparison of the Huber-LRT detector with existing ENF detectors.

where $T_{\text{Huber},(\cdot)}$ represents the order statistics of the test statistic values. By employing the median, the threshold value is less influenced by anomalous observations, making it a more reliable and robust choice for detecting deviations such as extreme values or outliers in the data.

3 EXPERIMENTAL EVALUATION

In this Section, the experimental evaluation is conducted to compare the proposed method with the state-of-the-art detectors developed in [15, 23]. A comprehensive description of the real-world audio recordings within the ENF-WHU dataset, followed by a detailed account of the preprocessing steps undertaken and the experimental results, is presented.

3.1 Dataset

The evaluation of the Huber-LRT detector utilizes the ENF-WHU dataset [15], which comprises 100 real-world audio recordings. These recordings are captured at a sampling rate of 44.1 KHz on the Wuhan University campus. Among these 100 recordings, 50 are placed under the folder labeled `H1` as they contain the ENF signal, while the remaining 50 recordings, characterized by severe corruption or the absence of the ENF signal, are placed in the `H0` folder.

3.2 Data Preprocessing

Prior to the evaluation process, a four-step pre-processing procedure summarized in Table 1 is applied to the audio recordings in the ENF-WHU dataset. Initially, the recordings are cropped into audio clips, with duration ranging from 5 to 10 seconds. Subsequently, the cropped audio clips undergo downsampling to 8 KHz using appropriate antialiasing filtering, followed by resampling at 400 Hz. Finally, a bandpass filter is implemented with the passband centered at the second harmonic of ENF (i.e., 100 Hz), and the cut-off frequencies are set at 99.9 Hz and 100.1 Hz. The transition bands have a width of 0.1 Hz. This pass-band choice aligns with the methodology outlined in [15], which centers around the second harmonic of ENF due to its robustness and prominence.

3.3 Experimental Results

In Figure 2, a comparative performance analysis of the proposed Huber-LRT detector against the existing methods [15, 23] is presented. Figure 2a displays the detection accuracy across different recording durations, spanning from 5 to 10 seconds. Notably, higher accuracy is achieved by the Huber-LRT detector, which is defined as the ratio of correctly detected instances to the total number of recordings. Notably, the Huber-LRT detector outperforms its competitors in terms of accuracy, which is defined as the ratio of correctly detected instances to the total number of recordings. More specifically, the Huber-LRT detector achieves an accuracy of 80% for 5-second audio durations and surpasses 90% for audio recordings lasting 7 seconds. However, a decline in accuracy is noticeable for all competing methods when dealing with 8-second audio segments. Nevertheless, the Huber-LRT detector maintains a satisfactory detection accuracy of 85%. In Figure 2b, the Receiver Operating Characteristic (ROC) curves for the proposed detector and its competitors are illustrated, accompanied by their corresponding Area Under the Curve (AUC). The AUC for the Huber-LRT detector is calculated to be 0.942, achieving a higher performance compared to other methods.

The improved performance of the Huber-LRT detector is attributed to the incorporation of robust statistical techniques in the ENF detection problem. By utilizing Huber regression for the estimation of the unknown parameters and assuming a mixture noise model, the ENF detection task is reinforced with greater strength. This assumption represents a significant departure from the LAD-LRT

detector [23], which centered around the assumption of a Laplacian noise model. The Huber-LRT provides a more efficient and robust method for modeling noise within the ENF-WHU dataset by combining elements of the Gaussian distribution in the middle and the Laplacian distribution in the tails. Consequently, the Huber-LRT, in conjunction with the previous methods [15, 23], contributes to a deeper understanding of ENF detection by leveraging the power of robust statistical methods.

4 CONCLUSION AND FUTURE WORK

Here, a novel ENF detector termed Huber-LRT has been proposed, incorporating robust statistical techniques into the ENF detection task. This innovative framework is built upon the assumption of a mixture noise model, combining Gaussian and Laplacian noise components to enhance ENF detection accuracy in short-length audio recordings. The effectiveness of this proposed ENF detector has been assessed through the accuracy and analysis of ROC curves with respect to AUC. Real-world benchmark data from the ENF-WHU dataset has been employed for this evaluation. The experimental findings have indicated a significant enhancement in ENF detection for short-length audio recordings, outperforming the performance of state-of-the-art techniques. Further research will consider the integration of other robust regression methods in the ENF detection task aiming at further improvements in detection accuracy.

ACKNOWLEDGMENTS

This research was supported by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the "2nd Call for H.F.R.I. Research Projects to support Faculty Members & Researchers" (Project Number: 3888).

REFERENCES

[1] D. Bykhovsky and A. Cohen. 2013. Electrical network frequency (ENF) maximum-likelihood estimation via a multitone harmonic model. *IEEE Transactions on Information Forensics and Security* 8, 5 (2013), 744–753.
[2] A. J. Cooper. 2009. An automated approach to the Electric Network Frequency (ENF) criterion: Theory and practice. *International Journal of Speech, Language & the Law* 16, 2 (2009), 193–218.
[3] S. Dash and N. C. Sahoo. 2022. Electric energy disaggregation via non-intrusive load monitoring: A state-of-the-art systematic review. *Electric Power Systems Research* 213 (2022), 108673.
[4] R. Garg, A. Hajj-Ahmad, and M. Wu. 2013. Geo-location estimation from electrical network frequency signals. In *Proceedings of the 2013 IEEE International Conference on Acoustics, Speech, and Signal Processing*. IEEE, Vancouver, Canada, 2862–2866.
[5] R. Garg, A. L. Varna, and M. Wu. 2011. "Seeing" ENF: natural time stamp for digital video via optical sensing and signal processing. In *Proceedings of the 19th ACM International Conference on Multimedia*. ACM, Scottsdale, AZ, 23–32.
[6] R. Garg, A. L. Varna, and M. Wu. 2012. Modeling and analysis of electric network frequency signal for timestamp verification. In *Proceedings of the 2012 IEEE International Workshop on Information Forensics and Security*. IEEE, Tenerife, Spain, 67–72.
[7] G.-O. Glentis. 2008. A fast algorithm for APES and Capon spectral estimation. *IEEE Transactions on Signal Processing* 56, 9 (2008), 4207–4220.
[8] G.-O. Glentis and A. Jakobsson. 2011. Efficient implementation of iterative adaptive approach spectral estimation techniques. *IEEE Transactions on Signal Processing* 59, 9 (2011), 4154–4167.
[9] C. Grigoras. 2007. Applications of ENF criterion in forensic audio, video, computer and telecommunication analysis. *Forensic Science International* 167, 2-3 (2007), 136–145.
[10] A. Hajj-Ahmad, R. Garg, and M. Wu. 2013. Spectrum combining for ENF signal estimation. *IEEE Signal Processing Letters* 20, 9 (2013), 885–888.
[11] H.-P. Hsu, Z.-R. Jiang, L.-Y. Li, T.-C. Tsai, C.-H. Hung, S.-C. Chang, S.-S. Wang, and S.-H. Fang. 2023. Detection of Audio Tampering Based on Electric Network Frequency Signal. *Sensors* 23, 16 (2023), 7029.
[12] G. Hua. 2018. Error analysis of forensic ENF matching. In *Proceedings of the 2018 IEEE International Workshop on Information Forensics and Security*. IEEE, Honk Kong, 1–7.
[13] G. Hua, G. Bi, and V. L. L. Thing. 2017. On practical issues of electric network frequency based audio forensics. *IEEE Access* 5 (2017), 20640–20651.
[14] G. Hua, J. Goh, and V. L. L. Thing. 2014. A dynamic matching algorithm for audio timestamp identification using the ENF criterion. *IEEE Transactions on Information Forensics and Security* 9, 7 (2014), 1045–1055.
[15] G. Hua, H. Liao, Q. Wang, H. Zhang, and D. Ye. 2020. Detection of electric network frequency in audio recordings–from theory to practical detectors. *IEEE Transactions on Information Forensics and Fecurity* 16 (2020), 236–248.
[16] G. Hua, H. Liao, H. Zhang, D. Ye, and J. Ma. 2021. Robust ENF estimation based on harmonic enhancement and maximum weight clique. *IEEE Transactions on Information Forensics and Security* 16 (2021), 3874–3887.
[17] P. J. Huber. 1992. Robust estimation of a location parameter. In *Breakthroughs in statistics: Methodology and distribution*. Springer, New York, NY, 492–518.
[18] G. Karantaidis and C. Kotropoulos. 2018. Assessing spectral estimation methods for electric network frequency extraction. In *Proceedings of the 22nd Pan-Hellenic Conference on Informatics*. ACM, Athens, Greece, 202–207.
[19] G. Karantaidis and C. Kotropoulos. 2019. Efficient Capon-based approach exploiting temporal windowing for electric network frequency estimation. In *Proceedings of the 2019 IEEE International Workshop on Machine Learning for Signal Processing*. IEEE, Pittsburgh, PA, USA, 1–6.
[20] G. Karantaidis and C. Kotropoulos. 2021. Blackman–Tukey spectral estimation and electric network frequency matching from power mains and speech recordings. *IET Signal Processing* 15, 6 (2021), 396–409.
[21] M. Kaselimi, E. Protopapadakis, A. Voulodimos, N. Doulamis, and A. Doulamis. 2022. Towards trustworthy energy disaggregation: A review of challenges, methods, and perspectives for non-intrusive load monitoring. *Sensors* 22, 15 (2022), 5872.
[22] S. M. Kay. 1993. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, Inc., Upper Saddle River, NJ.
[23] C. Korgialas and C. Kotropoulos. 2023. Electric Network Frequency Detection Using Least Absolute Deviations. In *Proceedings of the 2023 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, Rhodes, Greece, 1–5.
[24] H. Liao, G. Hua, and H. Zhang. 2021. ENF detection in audio recordings via multi-harmonic combining. *IEEE Signal Processing Letters* 28 (2021), 1808–1812.
[25] E. Ngharamike, L.-M. Ang, K. P. Seng, and M. Wang. 2023. ENF Based Digital Multimedia Forensics: Survey, Application, Challenges and Future Work. *IEEE Access* 11 (2023), 101241–101272. https://doi.org/10.1109/ACCESS.2023.3312181
[26] O. Ojowu, J. Karlsson, J. Li, and Y. Liu. 2012. ENF extraction from digital recordings using adaptive techniques and frequency tracking. *IEEE Transactions on Information Forensics and Security* 7, 4 (2012), 1330–1338.
[27] P. M. G. I. Reis, J. P. C. L. da Costa, R. K. Miranda, and G. Del Galdo. 2016. Audio authentication using the kurtosis of ESPRIT based ENF estimates. In *Proceedings of the 2016 International Conference on Signal Processing and Communication Systems*. IEEE, Gold Coast, Australia, 1–6.
[28] P. M. G. I. Reis, J. P. C. L. da Costa, R. K. Miranda, and G. Del Galdo. 2016. ESPRIT-Hilbert-based audio tampering detection with SVM classifier for forensic analysis via electrical network frequency. *IEEE Transactions on Information Forensics and Security* 12, 4 (2016), 853–864.
[29] S. Vatansever, A. E. Dirik, and N. Memon. 2017. Detecting the presence of ENF signal in digital videos: A superpixel-based approach. *IEEE Signal Processing Letters* 24, 10 (2017), 1463–1467.
[30] A. Verma, A. Anwar, M. A. Mahmud, M. Ahmed, and A. Kouzani. 2021. A comprehensive review on the NILM algorithms for energy disaggregation. arXiv:1312.5602
[31] Z.-F. Wang, J. Wang, C.-Y. Zeng, Q.-S. Min, Y. Tian, and M.-Z. Zuo. 2018. Digital audio tampering detection based on ENF consistency. In *Proceedings of the 2018 International Conference on Wavelet Analysis and Pattern Recognition*. IEEE, Chengdu, China, 209–214.
[32] C.-W. Wong, A. Hajj-Ahmad, and M. Wu. 2018. Invisible geo-location signature in a single image. In *Proceedings of the 2018 IEEE International Conference on Acoustics, Speech, and Signal Processing*. IEEE, Calgary, AB, Canada, 1987–1991.
[33] W. Yao, J. Zhao, M. J. Till, S. You, Y. Liu, Y. Cui, and Y. Liu. 2017. Source location identification of distribution-level electric network frequency signals at multiple geographic scales. *IEEE Access* 5 (2017), 11166–11175.
[34] C. Zeng, S. Kong, Z. Wang, K. Li, and Y. Zhao. 2023. Digital Audio Tampering Detection Based on Deep Temporal–Spatial Features of Electrical Network Frequency. *Information* 14, 5 (2023), 253.
[35] L. Zheng, Y. Zhang, C. E. Lee, and V. L. L. Thing. 2017. Time-of-recording estimation for audio recordings. *Digital Investigation* 22 (2017), S115–S126.
[36] A. M. Zoubir, V. Koivunen, E. Ollila, and M. Muma. 2018. *Robust Statistics for Signal Processing*. Cambridge University Press, Cambridge, MA.