The ENF criterion: A literature study

Tom Baksteen

June 2, 2015

Contents

1 Introduction

The number of digital audio and video recordings has been increasing over the last few decades. And with that also tools to modify its content and meta-data like time of creation. Some edits seem unnoticeable and the need to verify the authenticity of digital video and audio emerges. In legal cases recordings are used to indicate a person's presence at a certain location at a certain time. One could forge an alibi by editing footage. For this reason the Electrical Network Frequency (ENF) Criterion has been developed to authenticate such recordings. This technique allows for detecting cuts in the audio (or video) and determining the time of recording under some circumstances. The frequency of the power grid is used as a unique fingerprint which could unintentionally be found in the audio track. The first article to address the possibility of validating audio is written bij Catalin Grigoras in 2003 [8]. This was published in 2005.

The key observation is that the ENF fluctuates around its mean in a semi-random way. In Europe the power grid runs on 50 Hz and the ENF is controlled to be within a band of 200 mHz around this mean value. Each power withdrawal in Europe effects the frequency, hence there are so many components which influence this ENF that it becomes a like a white noise around the 50 Hz.

Secondly, the ENF is nearly identical on each location of the same interconnected power grid. The generators, which produce the electricity, rotate at the frequency of approximately 50 Hz. They are synchronous machines, meaning that they operate at the same frequency all over the grid. Hence, the electricity produced at a certain moment has the same frequency on the entire grid.

These properties allow for a timestamping of some digital audio files anywhere in a power grid when there is an ENF database available for pattern matching. An other powerful application of the ENF criterion is to detect edits in the audio recording [8]. These may be visible as large frequency discontinuities.

2 The ENF signal

The ENF signal which is studied is the signal consisting of one value per second, which is the most prominent frequency of the voltage on the power grid during each second separately. In order to gain some understanding concerning the ENF signal it is important to explore the basic workings of the power grid.

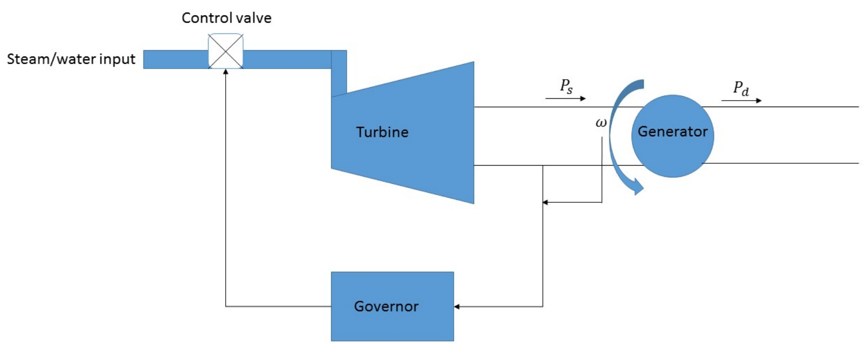

Figure 1: How the frequency is controlled

Figure 1 shows some of the dynamics of electricity generation. Turbines generate power which is fed into generator. The supplied power (P_s) makes the generator rotate so that kinetic energy is converted to electrical energy to meet the power demand of the grid (P_d). The spinning frequency of the generator is exactly the voltage frequency, the ENF.

The frequency is a constant 50 Hz whenever P_s equals P_d but since the power demand changes with every consumers light switch or use of electric machines, the demand P_d is far from constant. Hence the supply must instantaneously follow the demand. Since a change in power supply takes some time, and electricity can't wait, the needed power is drawn from the turbine's rotational energy causing it to slow down. So an increase in power demand results in a temporary decrease in ENF and vice versa [13]. The frequency may freely fluctuate within a band of 0.02 mHz around the mean value of 50. Outside this narrow band the primary control is activated meaning that a governor will control the turbine input to adjust P_s at will. In case of more severe frequency deviations, over 0.2 Hz, more drastic measures are taken like shutting down generators intentionally causing a power outage to prevent more serious damage from happening.

The energy market also plays a big role in the power supply to the generators. Each hour the right to sell power to generators is auctioned but the allocation of the required power amongst the generators has to be decided before the load appears. Therefore the load must be predicted in advance each hour. If an increase in load is expected, the turbine power will be increased in beforehand resulting in an increased frequency to anticipate the coming frequency drop.

Since there are many generators and turbines all over the connected grid, this applies to all generators. The power supply for each generator is managed relative to its mass. The generators are synchronous machines meaning that the frequency of all generators are, if all runs correctly, nearly identical.

As for the dynamics of the ENF there is a simple equation describing the relation between power and frequency. This excludes the control mechanisms however. Let ω denote the angular velocity of the generators, P_d and P_s be the power demand and power supply respectively. Furthermore we denote the cumulative kinetic energy of the rotating generators by H.

The following relation holds [13]:

$$\frac{d\omega}{dt} = \frac{1}{2H}(P_s - P_d) \tag{1}$$

It can be deduced that indeed the frequency increases whenever the supply exceeds the demand. The surplus is stored as kinetic energy in the rotation of the generators and a shortage is drawn from the rotation. The relation is approximately linear.

2.1 Obtaining the ENF

This chapter describes the process of estimating the ENF pattern from a given audio file. Although there are different techniques for this, the one treated here is the most commonly used technique. Cooper gives an overview of the overall ENF signal processing procedure [4].

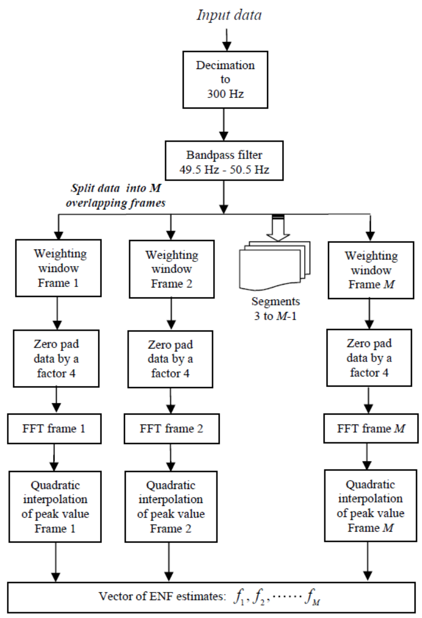

Figure 2: ENF processing procedure

As input for the procedure we have an audio recording. Most audio recordings are sampled at 44.1 kHz which is more than enough for the ENF extraction purposes. The Nyquist frequency, named after Harry Nyquist, is half the sampling rate of the recording [2]. All frequencies less than the Nyquist frequency can be detected without aliasing. Whenever audio is sampled at well more than 100 Hz, the ENF of 50 Hz can be detected in a straightforward way. In order to save memory and computation time Cooper downsamples the input signal to 300 Hz.

Secondly, a bandpass filter may be needed. A transformation to the frequency domain is applied and the filter preserves all possible ENF frequencies and attenuates frequencies from noise outside the filter band. At first it may seem unnecessary to filter out frequencies outside the 49.5 to 50.5 Hz band. After all, to determine the ENF value the frequency with the greatest Fourier coefficient has to be calculated. This is the location of the maximum of the signal in the frequency domain and we already know it to be within this band of 49.5-50.5 Hz.

The reason to apply a bandpass filter however is spectral leakage as depicted in Figure 3. The spectrum of a perfect 50 Hz signal is not one peak at 50 and zeros elsewhere but a sinc function with a maximum at 50 Hz and what is called spectral leakage. The image suggests that there is a nonzero component of e.g. 48.5 Hz but this is only a sidelobe belonging to the mainlobe of 50 Hz. There is no 48.5 Hz sinusoidal contribution to the signal at all.

In the same way, some background noise in the recording which *is* 48.5 Hz has spectral leakage which will affect the peak value at 50 Hz and also the ENF estimate. A bandpass filter will decrease this undesired effect.

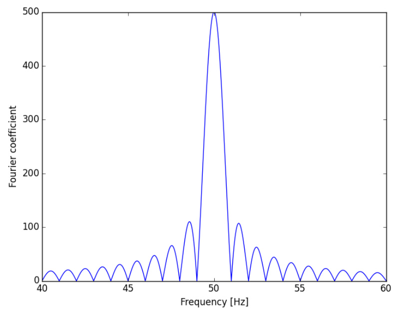

Figure 3: Spectrum of a 50 Hz signal: Spectral leakage

2.1.1 Short Time Fourier Transform and zero padding

The Short Time Fourier Transform (STFT) is implemented as a Fast Fourier Transform (FFT) with a sliding window. It is a transformation from the time domain to the frequency domain and provides per frequency bin the corresponding spectral magnitude.

Given our (discrete) signal $s[n]$ and a chosen symmetric window function $w[n]$ of length N, the STFT of s are the complex coefficients:

$$X(m,k) = \sum_n x[n]w[n-m]e^{-j2\pi kn/n}$$

Here m determines the position of the window function and hence the time segment to be transformed. The time step, which is the time resolution, is set to one second in most literature. k is the frequency index corresponding to the frequency kN.

The uncertainty principle states that it is impossible to improve both the time resolution and the frequency resolution of a signal beyond a certain point [2]. There are however artificial ways to surpass this principle. Zero padding is often called frequency domain interpolation [12]. It does not actually increase the frequency resolution but interpolates additional frequency bins. Zero padding basically involves adding a certain number of zeros to the windowed signal segment. This artificially increases the length of the segment, allowing the FFT to return more frequency magnitudes.

The FFT of 1 second signal without zero padding returns values for frequency bins with 1 Hz resolution. When zero padding is applied to increase the length of a segment from 1 second to 10 seconds, which is adding 9 seconds worth of zeros, the frequency resolution is enhanced to 0.1 Hz. For the purpose of determining the time of recording an accuracy of at least 1 mHz is required. This implies a zero padding factor of 1000 is needed. This is however not feasible due to memory and computation time. Some other technique should be used additionally.

More accurate ENF estimation can be obtained by quadratic interpolation over the most prominent frequency bucket and its direct neighbors. This technique is computationally light and uses little memory.

2.1.2 Quadratic Interpolation

Rife and Boorstyn [12] show that when using a rectangular window, one can obtain an exact spectral peak estimator as the maximum likelihood estimator for a single sinusoid with additive white Gaussian noise by zero padding to infinity. The maximum likelihood estimator will asymptotically achieve the Cramer-Rao lower bound on the variance of the peak frequency estimator.

Deriving the STFT with too much zero padding becomes computationally expensive and requires a lots of memory. So the commonly used spectral peak estimator is one which requires less memory, the Quadratically Interpolated FFT (QIFFT).

Let m denote the frequency bin with the highest FFT value. The largest frequency component is then assumed to be in in one of the bins $m-1, m$ and $m+1$. To determine the estimated ENF more accurately than stating that it is somewhere in this frequency bin, quadratic interpolation is used on the three FFT values.

Given $\mathrm{FFT}(m-1) = \alpha, \quad \mathrm{FFT}(m) = \beta, \quad \mathrm{FFT}(m+1) = \gamma$ the relative position of the interpolated top can be calculated as:

$$p = \frac{1}{2} \cdot \frac{\alpha - \gamma}{\alpha - 2\beta + \gamma}$$

Meaning that the best estimation of the ENF is located at 'bin' $m + p$.

2.2 Noise

In the study of ENF noise is considered to be any sound other than the 50 Hz signal coming from the power grid. Since all relevant recordings do not intend to record the ENF the sound to noise ratio is typically very low, which can result in bad ENF estimations. It is said that noise is the biggest limitation of the ENF criterion [4]. Hence, noise reduction is the main topic in ENF literature. Many techniques have been proposed for noise reduction like a threshold dependent median filtering [4], using multiple ENF harmonics [3] or autocorrelation models [6], and increasing the STFT window overlap.

To illustrate the effect of different STFT overlap the following three figures show the ENF signal estimation of one recording using different STFT overlaps.

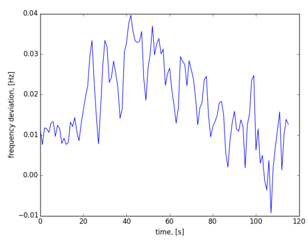

Figure 4: 4 sec overlap

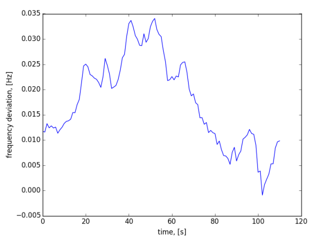

Figure 5: 9 sec overlap

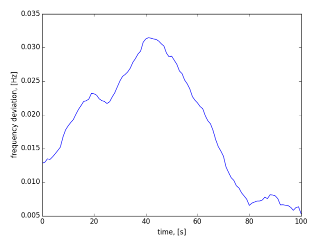

Figure 6: 19 sec overlap

All these images show an ENF estimation with 1 second resolution, as is the convention.

In the first image the ENF estimation using a window size of 5 seconds is shown, that is an overlap of 4 seconds. The pattern is much more jagged than what the real ENF without noise would produce.

One property of the ENF is that the frequency value from 1 second to the next will not jump more than 40 mHz. This is due to the inertia of the generators, meaning that they can't slow down that fast in 1 second. So Figure 4 does not preview an accurate ENF signal. To smoothen this pattern Cooper suggests to increase the window overlap.

Figure 5 shows the ENF estimation using a 9 second overlap in segments. As a result the frequencies appear more plausible than before and a less jagged signal is plotted. One could even further increase the overlap to 19 seconds to get the result displayed in Figure 6. No article has proposed a convention concerning the overlap yet, neither has there been claims about the overlap factor producing the best results. 1 to 14 seconds overlap have all been used, where the latter works better for visual matching with the true ENF signal.

3 ENF database matching

For the purpose of timestamping it is essential to set up a database of ENF values with their corresponding times. This is normally done by using a step down voltage directly connected to the power grid and either feeding the signal to a computer sound card which runs the FFT or by counting zero crossings.

The zero crossings method treats the signal as sinusoidal, and calculates the time τ between two consecutive zero crossings. The instantaneous frequency is then $f = \frac{1}{2\tau}$. The times of the zero crossings are determined by linear interpolation between consecutive samples that differ in sign. The frequency of one second of signal is the average of all frequency estimations obtained in that second. Such databases were already existing before the application of the ENF criterion since the power grid frequency is continuously being monitored. Note that the location where the database is created is irrelevant as long as it is on the same grid as the recording that is to be timestamped.

Mains powered audio recordings of a few minutes are mostly sufficient to accurately timestamp it. Under two minutes the recording may be too short, resulting in the loss of uniqueness of the ENF extract. A wrong database match could be found. [9] Hence, the length of the recording is another big issue along with the noise.

The database matching technique all authors use is straightforward: The estimated ENF from recording of n seconds is compared to the first n seconds of database values, then to the segment shifted one second forward etc. The database segment which is the closest match to the extracted ENF signal indicates the time of recording.

Now the question rises how the best match should be defined. The literature provides two norms: the Cross Correlation (CC) on the other hand and the Mean Squared Error (MSE) on the other hand. The CC of the two ENF vectors x and y is defined as $CC = \frac{\sum x \cdot y}{N \cdot \sigma_x \sigma_y}$. To find a match the greatest correlation coefficient in the database has to be found. This is more or less equivalent to finding a minimum MSE where $MSE = \frac{\sum (x-y)^2}{N}$. Here x can be treated as the extracted ENF signal and y the selected database segment.

Many articles have compared the performance of Cross Correlation (CC) with Mean Squared Error (MSE). There have been claims that maximizing the CC matching should be preferred over minimizing the MSE since it would increase the chance of accurately timetamping recordings [9]. Others claim the MSE provides better results, when the two vectors are zero mean [4].

The MSE distance norm without subtracting the vector mean may give bad results since some recording devices have a frequency offset due to its hardware [11] [10]. When the vectors are both zero mean the frequency offset does not play any role in the matching process.

An article about weather forecasts by Barnston [1] shows a direct relation between the CC and the MSE. First we define ZMMSE to be the zero mean MSE, then the following relation holds:

$$\begin{aligned} ZMMSE &= \frac{\sum (x-y)^2}{N} \\ &= \frac{\sum x^2 + \sum y^2 - 2\sum x \cdot y}{N} \\ &= \sigma_x^2 + \sigma_y^2 - 2CC \cdot \sigma_x \sigma_y \\ &= 2\sigma^2(1 - CC) \end{aligned} \tag{2}$$

Here the assumption is made that the standard deviations of the extracted ENF is equal to its match in the database, which is a fair assumption.

The relation between the correlation coefficient and the zero-mean mean squared error is one-to-one and linear. Therefor it is unlikely that one would give better results than the other in terms of precision. Hence it is advised to use the ZMMSE approach to minimize the computational work.

Minimizing $\sum (x-y)^2$ over all database segments seems to be an efficient matching algorithm. (Note that dividing by the constant vector length is superfluous). But now the question rises what ZMMSE score should be considered a convincing match. Two matching score criteria will be discussed next.

3.1 Interdatabase error

The first technique is developed by Huijbregtse and Geradts [9]. They had an ENF database of 1.5 years of frequency values. From this data set two randomly picked non-overlapping segments of 600 ENF values (10 minutes) are compared to each other. This is repeated one million times and the matching scores (CC or MSE) are stored and plotted.

The smallest root mean squared error between two random segments was found to be 4 mHz. Based on this observation they state that for the 10 minutes ENF vector of interest, the root mean squared error with respect to a database segment should be "well below" 4 mHz. Whenever this is not the case, the strength of the match is not good since a better match has been found between 2 random disjoint samples taken from that database. A false match is not unlikely.

3.2 Error distribution

An other technique is described by Cooper [4]. The errors of the ENF extract with respect to all the database segments form a distribution. In certain experiments they found the error is normally distributed and for some other the distribution was skewed. In case of the skewed distribution it can be transformed to a normal distribution. Now the point of smallest error is located a certain number of standard errors under the mean error. This is called the z-score and is used to quantify the strength of the match. Furthermore, plotting all errors in one graph also graphically shows how small the error is compared to the other errors. When a few of the smallest errors are relatively close to each other, the match is also questionable.

4 ENF signal from video

As said earlier, some audio recordings are contaminated with too much noise to make ENF extraction impossible for timestamping purposes. In some cases there is only video available without sound, like when security tapes are brought up as evidence in legal cases. The authors of "Seeing ENF" [5] examined the possibility of ENF estimation from video frames. Obviously there is no 50 Hz component present in the sound track so there has to be some mechanism visually expressing the network frequency. One such observation is that fluorescent light blinks with twice the ENF. Applying current at the nominal ENF of 50 Hz, its polarity changes at twice that rate and the current path is switched on and off at about 100 Hz. For example, as a stationary camera records a white wall with fluorescent light in the room the frames will all have different light intensity depending on the phase of the current. Our eyes process an unchanging white wall and can not register 100 Hz flickering but the camera does.

There is however one problem with the frames intensity registration. That is the low sampling rate of most video equipment. A frame rate of 30 fps is a common setting and as stated before it is generally only possible to detect frequency components of up to half the frame rate, which would be 15 Hz in this case. Fortunately the aliasing effect allows a trace back of the 100 Hz signal due to symmetry of the frequency spectrum around $f_s/2$ and duplications of the spectrum after each f_s Hz. This duplication of the complete Fourier spectrum is due to the fact that $sin(\omega t)$ is indistinguishable from $sin((\omega + f_s)t$ when sampled with f_s Hz.

To illustrate the effect, a camera running at 29,97 Hz will have a spectrum symmetric around 14,985 Hz and repeating each 29,97 Hz. Hence, the 100 Hz component will be detected at the frequency of 10,09 Hz where it is shifted with $3f_s$. 10.09 Hz is called the alias of 100 Hz. The second harmonic of the fluorescent signal is located at 200 Hz. Subtracting $7f_s$ gives -9.79 Hz. The symmetry property then shows the 200 Hz harmonic ENF at 9.79 Hz.

It must be noted however that the ENF signal consisting of these second harmonic aliases has a doubled bandwidth, meaning that each ENF deviation from the mean 50 Hz will result in a doubled deviation from the 9.79 Hz value. Also, these values are mirrored since the positive frequencies around 9.79 are actually symmetries of the negative frequencies. For example, an ENF of 49.98 Hz will be visible in aliases at 10.07 from the fundamental 100 Hz frequency but also at 9.80 from the second harmonic. In the same way, if the third harmonic is also present in the fluorescent light, the 50 Hz ENF will have its alias at 0.30 Hz.

This makes ENF estimation from video still possible. It is very hard to get a solid estimate because the frame rate is very low and different harmonics can be projected on the same frequencies for certain frame rates. The aliasing effect on 30 Hz video recording device will cause the 100 Hz fundamental and the second harmonic to be both projected onto 10 Hz, making the ENF extraction very complicated if not impossible.

5 Location estimation of a recording

One most interesting topic in the study of the ENF criterion is whether it is possible to also accurately estimate the location of a recording.

By most researchers this question is answered with no. The ENF is assumed to be the same at each location of an interconnected power grid. Only one article has been written as an attempt to be able to locate a recording [7]. Some small changes in load may only have a local effect on the frequency. Anticipating the ENF patterns are more similar for locations nearby than for locations at greater distance, the authors empirically found that the frequencies do not exactly coincide on different locations. The measurements done in that research were in the United States. Five ENF databases at 200 to 1200 kilometers apart were created and a spread in ENF values of 1 to 2 mHz was observed.

One goal was to gain understanding about the location dependence of the ENF signal and increase timestamping by comparing a recording with a nearby established database. The other part of the study consist of doing a location estimation on recordings. They call the used method Half-plane intersection.

Let X be the unknown location of some audio recording and A, B, ... are database recording locations. X is assumed to be closer to A than to B when the signal of X is stronger correlated to the ENF pattern from database A than it is to B. X should then be located in the half-plane consisting of all points closer to A.

This is repeated for each pair of database locations and X is likely located in the intersection of those half-planes.

The article however does not provide any theoretical insight why this would work even though it appears to give solid results using their (noiseless) data.

5.1 Theoretical analysis

A theoretical ground for the statements made about estimation of geo-location can be found in literature about power grids.

The famous Telegrapher equations describe the voltage and current on an electrical transmission line over time and distance. The coupled first order PDE's for a lossy transmission line in the time domain are given by:

$$\begin{aligned} \frac{\partial}{\partial x}V(x,t) &= -L\frac{\partial}{\partial t}I(x,t) - RI(x,t) \\ \frac{\partial}{\partial x}I(x,t) &= -C\frac{\partial}{\partial t}V(x,t) - GV(x,t) \end{aligned} \tag{3}$$

Alternatively, equations (3) can be formulated as two decoupled second order equations:

$$\begin{aligned} \frac{\partial^2}{\partial x^2}V(x,t) &= LC\frac{\partial^2}{\partial t^2}V(x,t) + (RC+GL)\frac{\partial}{\partial t}V(x,t) + GRV(x,t) \\ \frac{\partial^2}{\partial x^2}I(x,t) &= LC\frac{\partial^2}{\partial t^2}I(x,t) + (RC+GL)\frac{\partial}{\partial t}I(x,t) + GRI(x,t) \end{aligned} \tag{4}$$

For negligible isolation material conductance G, this reduces to:

$$
\begin{aligned}
\frac{\partial^2}{\partial x^2}V(x,t) &= LC\frac{\partial^2}{\partial t^2}V(x,t) + RC\frac{\partial}{\partial t}V(x,t) \\
\frac{\partial^2}{\partial x^2}I(x,t) &= LC\frac{\partial^2}{\partial t^2}I(x,t) + RC\frac{\partial}{\partial t}I(x,t)
\end{aligned} \tag{5}
$$

Which are wave equations with an extra first order term causing the signals to decay over time and distance.

Starting with the Telegrapher equations in the frequency domain:

$$
\begin{aligned}
\frac{\partial}{\partial x}V &= -(R + j\omega L)I \\
\frac{\partial}{\partial x}I &= -(G + j\omega C)V
\end{aligned} \tag{6}
$$

can be written as:

$$
\begin{aligned}
\frac{\partial^2}{\partial x^2}V &= \gamma^2 V \\
\frac{\partial^2}{\partial x^2}I &= \gamma^2 I
\end{aligned} \tag{7}
$$

where $\gamma = \sqrt{(R + j\omega L)(G + j\omega C)}$ is the propagation constant and ω the angular frequency. Assuming relatively small losses R and G compared to the frequency we obtain an approximation for $V(x_2, t+\tau)$, given an input pulse $V(x_1, t)$:

$$
V(x_2, t+\tau) \approx V(x_1, t)e^{-\frac{\sqrt{LC}}{2}\left(\frac{R}{L}+\frac{G}{C}\right)(x_2-x_1)} \tag{8}
$$

τ is the travel time of the electric signal from x_1 to x_2. It holds that $\tau = \sqrt{LC}x$.

Equation (8) can be manipulated to investigate the behavior of the cross correlation of a signal with itself at a distance.

Multiplying both sides with $V(x_1, t)$ and time averaging results in:

$$
\langle V(x_2, t+\tau)V(x_1, t)\rangle \approx \left\langle (V(x_1, t))^2 \right\rangle e^{-\frac{\sqrt{LC}}{2}\left(\frac{R}{L}+\frac{G}{C}\right)(x_2-x_1)} \tag{9}
$$

This shows that the correlation coefficient of the power grid signals measured at two locations is maximal when $x_1 = x_2$ and wears off exponentially with the increase of the distance.

A second observation, supporting the assumption of a location dependent ENF signal is that when the load changes locally, the generator responsible for that area will be burdened with the consequences, a change in ENF. Let's say the load increases. This reduces the generator speed. This generator will at that short moment spin slower (or faster) than the other generators and becomes a load to the other generators and all generators frequency will be leveled again. Since signals travel at finite speed and generators large inertia prevent them from instantaneously changing in rotational speed the frequency drop caused by the load increase will gradually spread through the grid. With accurate measuring instruments these different ENF signals are observable.

6 Conclusion

The ENF criterion is a fairly new field of research with only ten years of history. On many different topics within the ENF study, the knowledge is still increasing and there are many topics left to research and optimize. Among those noise filtering may be the most important in order to increase timestamping accuracy. Fortunately there is already much knowledge about digital filters and such.

There is an ongoing debate whether to use the Cross Correlation or the Mean Squared Error as a database matching norm. But other literature, not concerning ENF, has shown that these two are equivalent when comparing two zero mean vectors, so that should close that chapter. In general, most of the ENF literature does not provide any rigorous mathematical proofs and most observations are solely empirical. It is not yet clear which STFT overlap factor should be maintained and what padding factors are acceptable.

Since two years the ENF studies also aim at validating video recording by using fluorescent light properties. Duo to the low framerate of most camera's and the aliasing effect of multiple harmonics ENF estimation from video is difficult. The research done on this topic is still very limited though and much improvement is still possible. The fact that video frames are usually written from the top down one row at a time has not been used yet. This may improve ENF estimation significantly.

The article about estimating the geo-location of audio recordings also opened up a very interesting possibility that certain regions of the grid can be ruled out when trying to determine the location of recording. The findings from this research combined with the knowledge about signal propagation may suggest a more direct relation between the correlation of the ENF signals from two different locations and the distance between them. From a mathematical point of view there is a lot of improvement possible on that topic.

References

[1] Anthony G Barnston. Correspondence among the correlation, rmse, and heidke forecast verification measures; refinement of the heidke score. *Weather and Forecasting*, 7(4):699–709, 1992.

[2] Math H Bollen and Irene Gu. *Signal processing of power quality disturbances*, volume 30. John Wiley & Sons, 2006.

[3] D. Bykhovsky and A. Cohen. Electrical network frequency (enf) maximum-likelihood estimation via a multitone harmonic model. *Information Forensics and Security, IEEE Transactions on*, 8(5):744–753, May 2013.

[4] Alan J Cooper. An automated approach to the electric network frequency (enf) criterion-theory and practice. *International Journal of Speech, Language and the Law*, 16(2):193–218, 2009.

[5] R. Garg, A.L. Varna, A. Hajj-Ahmad, and Min Wu. "seeing" enf: Power-signature-based timestamp for digital multimedia via optical sensing and signal processing. *Information Forensics and Security, IEEE Transactions on*, 8(9):1417–1432, Sept 2013.

[6] R. Garg, A.L. Varna, and Min Wu. Modeling and analysis of electric network frequency signal for timestamp verification. In *Information Forensics and Security (WIFS), 2012 IEEE International Workshop on*, pages 67–72, Dec 2012.

[7] Ravi Garg, Adi Hajj-Ahmad, and Min Wu. Geo-location estimation from electrical network frequency signals. In *ICASSP*, pages 2862–2866, 2013.

[8] Catalin Grigoras. Digital audio recording analysis–the electric network frequency criterion. *International Journal of Speech Language and the Law*, 12(1):63–76, 2005.

[9] Maarten Huijbregtse and Zeno Geradts. Using the enf criterion for determining the time of recording of short digital audio recordings. In *Computational Forensics*, pages 116–124. Springer, 2009.

[10] Mateusz Kajstura, Agata Trawinska, and Jacek Hebenstreit. Application of the electrical network frequency (enf) criterion: A case of a digital recording. *Forensic science international*, 155(2):165–171, 2005.

[11] Alex Kantardjiev. *Determining the recording time of digital media by using the electric network frequency*. PhD thesis, Uppsala Universitet, 2011.

[12] D. Rife and R.R. Boorstyn. Single tone parameter estimation from discrete-time observations. *Information Theory, IEEE Transactions on*, 20(5):591–598, Sep 1974.

[13] Birron Mathew Weedy, Brian John Cory, N Jenkins, JB Ekanayake, and G Strbac. *Electric power systems*. John Wiley & Sons, 2012.