

# Digital Audio Recording Analysis

## The Electric Network Frequency Criterion

Catalin Grigoras, Ph.D, *IAFPA*

### INTRODUCTION

Over the last years we saw a significant increase in the number of attempts to use digital audio evidence in every sector of litigation and criminal justice. Clearly, many attorneys are facing the technical challenges of dealing with digital evidence and the forensic audio examiner has to analyze a digital audio recording (DAR), to authenticate its originality and content, and to write a forensic report.

### FORENSIC AUDIO ASPECTS

The aim of forensic audio tape analysis is to establish whether a recording on a tape is an original recording or a copy and to find out whether a recording was made on a given recorder (authenticate the recording). When electronic equipment is used to record a conversation on tape, it captures not only the intended speech but also the idiosyncrasies and characteristics of the recorder itself. Tape recorders leave start, stop and pause signatures coming from record and erase heads. Koenig [1], Pellicano [2], Dean [3], Molero [4] and others have shown that by combining waveform, spectral components, spectrograms and magnetic patterns analysis, the forensic audio examiner in most cases can establish the originality of a recording and/or authenticate the recording.

We live in a digital world and deal with a lot of digital evidence. When we analyse digital audio recordings, in most situations no idiosyncrasies left behind by the record head and/or erase head of the recorder can be found. Therefore a new approach is necessary.

Digital audio recordings can be classified in several ways (eg. compressed vs non-compressed, long-play vs short-play, digital audio tape vs memory or hard-disk, etc). The criterion I propose to classify DAR is:

- with 50/60 Hz frequency (and harmonics),
- without 50/60 Hz frequency.

When digital equipment (computer, DAT, etc.) is used to record a conversation, it captures not only the intended speech but also the 50/60 Hz electric network frequency (ENF), if the recording equipment is so powered and is lacking an ideal voltage regulator. Figure 1 displays the waveform and spectrogram of a real 50 Hz signal.

In a real electric network the ENF is not fixed at precisely 50 Hz. Over time frequency variations may occur because of differences between produced and consumed power. Although

the waveform and FFT can be used to examine the signal's periodicity on short time windows, the most relevant information on variation over time can be derived from the spectrogram where we can see the history of all peak values in time, representing the variation of the ENF.

In order to acquire the signal and to analyze its frequency, I used DCLive Forensics. There is no periodicity in the variation of the ENF and there is a random fluctuation of the ENF around 50 Hz. The power grid can have an instantaneous frequency variation of up to +/-0.6 Hz for up to a 10 seconds interval. When measured over an 8 or even 24-hour interval, the frequency tolerance is much tighter.

At one moment, the mathematical form of the ENF can be written like this:

$$f = [50 \pm \Delta f] \cdot Hz$$

where  $\Delta f$  represents the deviation between the instantaneous frequency and the set point frequency.

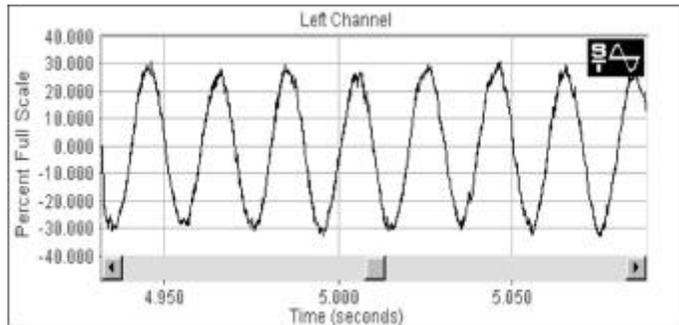
Due to the electromagnetic wave propagation, the entire electric network formed by the interconnected systems, including all the sources and loads, will carry the same frequency. Therefore, in an electric network we will find the same frequency through all the loads (houses, offices, labs, etc.). From a forensic point of view, frequency is the network's parameter we are interested to analyse.

According to the Union for Power Production and Transport Coordination (UCPTE) recommendation [7], under normal conditions the ENF is maintained within strict limits and the set point frequency is around 50 Hz. There are three types of operating conditions, based on  $\Delta f$ :

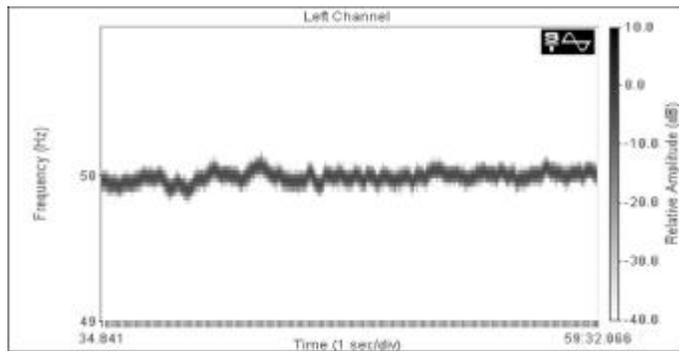
If  $\Delta f \leq 50$  mHz, the conditions are considered as normal, if  $50$  mHz  $< \Delta f \leq 150$  mHz operating conditions are deemed to be impaired, but with no major risk, and if  $\Delta f > 150$  mHz operating conditions are deemed to be severely impaired and there are significant risks of malfunction of the electric network.

From an engineering instrumentation point of view, all electric and electronic systems contain noise sources. To analyse the ENF it is recommended to use only high quality sampling devices with very low levels of noises and distortions. Based on my tests and results, a sound card with a

signal to noise ratio less than -94 dB and total harmonic distortions less than 0.003 % will offer you the possibility to analyse ENF without relevant errors and distortions.



(a) Waveform



(b) Spectrogram

Fig.1. The waveform and spectrogram of a real 50 Hz signal

The signal’s downsampling operation to 120 Hz is combined with a fft-band pass filter with the cut frequencies below 49 Hz and above 51 Hz. This filter eliminates the other signals on 0 – 60 Hz band and does not modify the ENF component which is around 50 Hz. For this purposes I used the “Change Sample Rate / Resolution” function in DCLive Forensics

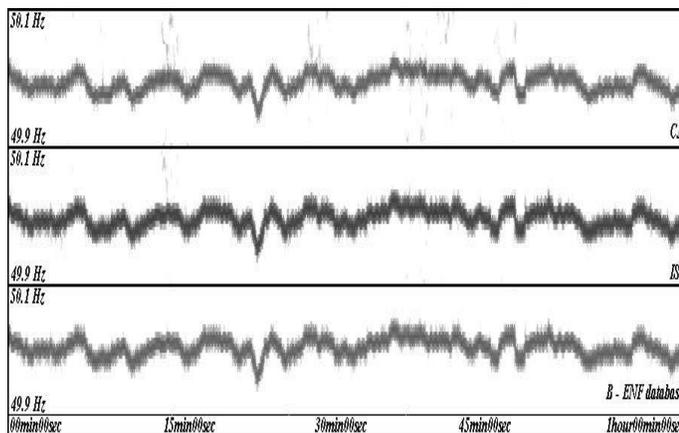


Figure 2. Spectrograms comparison

The tests I’ve made on different locations of the same electric network (in the lab, the same building, the same town and in three different cities all connected to the same electric network) showed that for all consumers, at any moment in time, the ENF value is the same. In the following figures are presented the most relevant tests, including system configuration and obtained results. In order to record and analyse the audio signal I used DCLive Forensics software. For three simultaneous recordings in the same electric network, the obtained results have a great degree of correlation and an example is presented in figure 2. Figure 2 shows that at any one moment in time, all the loads of the electric network experience the same ENF.

### A CASE WORK EXAMPLE

In a recent forensic case I was asked to analyse a digital recording brought to the court as evidence, to establish the authenticity of the recording and its originality. The evidence was an audio file on a CD, WAV PCM format, 16 bit, sampled at 8 KHz, containing more than three hours of conversation between two persons (speaker A and B) and it was made using a recording system installed in a room in an office building. I found that the system allows the audio signal to be picked-up from the room with a small electret microphone, amplified and saved in a WAV PCM format (no compression), 16 bit, 8 KHz, on the hard-disk of a notebook. Speaker A, who had previously been recorded by speaker B with this system, claimed that the recording was altered and that the claimed date of recording was not correct; he indicated another date. Information on both dates was given to the investigator.

I established the period of time on which to verify the ENF and tested whether there were any differences in the frequency variations between the ENF reference monitoring system and the frequency variations measured in the company’s building. I asked the electricity company if there was any break in the network over the last couple of months and I checked whether the building used it’s own private power supply or was connected to the network. I started by comparing the evidence with my ENF database to see if there was any match.

I analysed the audio file by using DCLive Forensics software, down sampled the evidence to 120 Hz (with “Change Sample Rate / Resolution” function), band-pass filtered the audio with cut frequencies around 49 and 51 Hz (with “Band Pass Filter” having the following settings: Low Freq = 49 Hz, High Freq = 51 Hz, Filter Slope = 24 dB/Octave, Butterworth), computed the spectrogram with a 4096 points FFT, made a vertical zoom around 50 Hz and compared with the ENF database. The resulting 2D-spectrograms are presented in figures 3, 4 and 5.

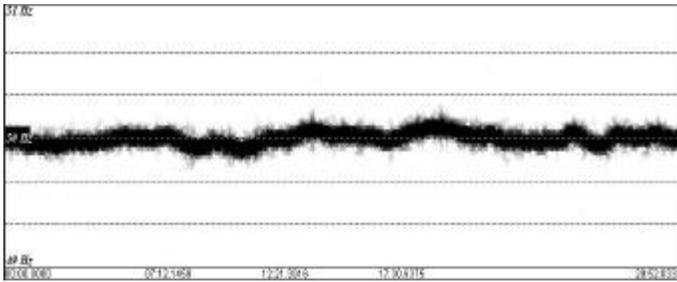


Figure 3. Real case analysis – ENF spectrogram claimed by speaker B

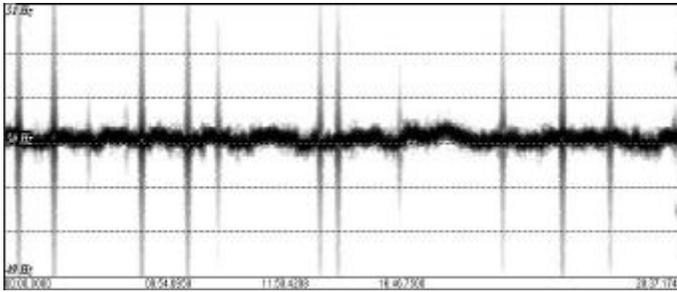


Figure 4. Real case analysis – Evidence ENF

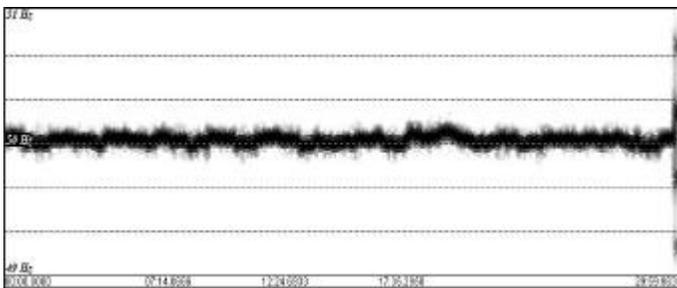


Figure 5. Real case analysis – ENF spectrogram claimed by speaker A

The conclusion is that the evidence had not been recorded on the date claimed by speaker B and there were matches with the date and time claimed by speaker A. I also checked the evidence’s ENF continuity and I discovered more than ten major discontinuities.

I asked speaker A to listen to the evidence, to indicate in the transcription where there are any irregularities and to write down what the original wording was. He indicated more than twenty deletions of words and expressions. The discontinuities in ENF that I had discovered matched his indications.

APPLICABILITY OF THE ENF CRITERION

The space necessary to save the ENF reference as WAV PCM, 16 bit, 120 Hz, is presented in table 1.

In order to save the reference over one month less than 630 MB is needed which can be saved on a CD and for one year this amounts to twelve CDs. It will be the examiner’s choice to save and to configure the entire reference database for fast access.

Table 1. HDD Space over maximum one day

Hours	Seconds	Capacity [MB]
1 hour	3600	0.843
8 hours	28800	6.744
16 hours	57600	13.488
24 hours	86400	20.232
1 week	604800	141.624
1 month	2678400	627.192
1 year	31536000	7384.68

The entire work to verify or identify the real date and time when a DAR was created needs hours of searching and to optimise it I propose an automatic system like the one presented in figure 6.

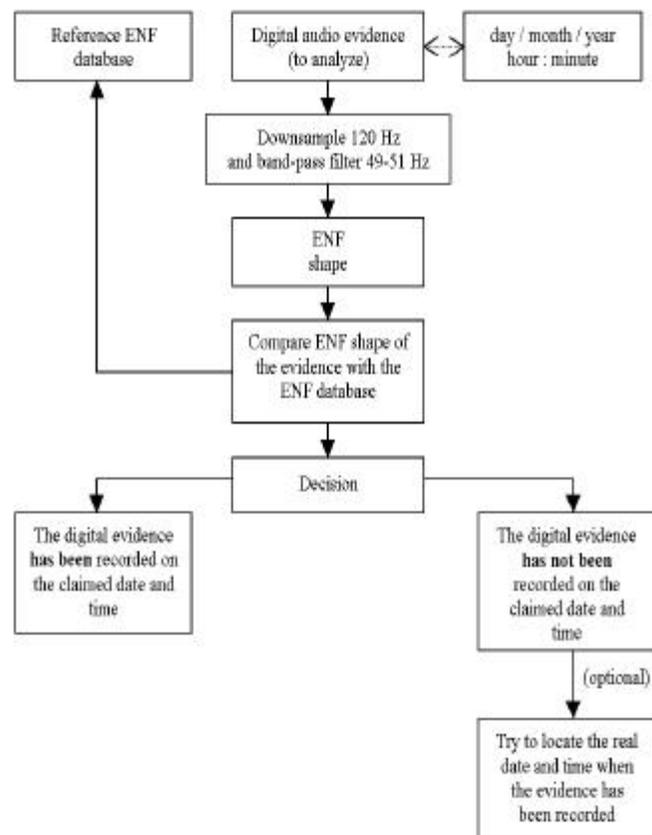


Figure 6. An automatic system to identify the date and time when a digital recording was created

As can be seen in figure 6, the idea in DAR authentication, based on the presented method, is to verify or try to locate the date and time when a DAR recording has been created and to search for modifications and deletions, to identify their length and to approximate how many samples have been deleted and how the audio segments are linked.

As an example of a DAT-recorder available to the public I analysed the Sony DAT Walkman TCD-D100 and its

recording particularities. I found that if recordings are made with this recorder powered from the electric network (through its own adapter), they will contain traces of the ENF and the methodology to analyse them as described above can be used. On digital evidence, the ENF varies around 50 Hz, but on analog tapes, because of wow and flutter, the mean value may not be around this value. Also if the play speed is faster than the recording speed, all the frequencies will be shifted up, and if played at a slow speed they will be shifted down and the measured ENF will need to be compensated for this shift.

On analog audio tape recordings, to use the ENF criterion, two problems remain a challenge:

- in the frequency domain, the vertical realignment on spectrogram, same as the horizontal compression of the spectrum (compensating speed differences),
- the horizontal adjustment in time domain, which means a stretch.

### CONCLUSIONS

The Electric Network Frequency Criterion is a tool that can be used to analyse digital audio recordings, checking their integrity and verifying or identifying the moments in time when a digital recording was created, by using a reference frequency database built in a laboratory or obtained from the electrical network company.

Particular attention should be given to the quality of the reference that is critical in allowing an accurate comparison, as well as to the entire methodology of analysis and evaluation of the results.

I recommend to use the ENF-Criterion in conjunction with other techniques and methods (e.g. impulse response of a room [5] and IT investigations) that must be improved and inserted into a standard procedure or guidelines.

By using the Electric Network Frequency Criterion to analyse an evidential recording the forensic audio examiner can transform himself into forensic audio “archeologist”, able to dig deep into audio archives for non-audible information.

### ACKNOWLEDGMENT

The author would like to thank Jos Bouten (Netherlands Forensic Institute), Peter French (JP French Associates, UK) and Alan Cooper (Metropolitan Police, London, UK) for helpful suggestions and comments on the paper.

I appreciate the support of my Prof. Lucian Ionescu (Director of National Institute of Forensic Expertise, Romania) and Prof. Brandusa Pantelimon (Politehnica University of Bucharest).

Finally, I would like to appreciate the very useful discussions I have had with Francis Nolan (Linguistics Department, Cambridge University), Olaf Köster (BKA, Germany), Didier Meuwly (Forensic Science Service, UK), Tom Owen (The New York Institute of Forensic Audio, USA), Agata Trawinska and Mateusz Kajstura (Institute of Forensic

Research in Cracow) and Curtis Crowe, Enhanced Audio, Inc. This manuscript could not have been completed without their help.

### REFERENCES

- [1] B.E. Koenig, “Authentication of Forensic Audio Recordings”, *J. Audio Eng. Soc.*, Vol. 38, No.1/2, 1990 January-February
- [2] A.J. Pellicano, “Tape Recordings as Evidence”, *California Lawyer*, October 1990
- [3] D.J. Dean, “The Relevance of Replay Transients in the Forensic Examination of Analogue Magnetic Tape Recordings”, in *Publication No.16/1991*, Scientific Research and Development Branch, Home Office, British Government, London
- [4] S. Molero, “Establishment of the Individual Characteristics of Magnetic Recording Systems for Identification Purposes”, *Problems of Forensic Sciences, Second EAFS meeting, Volume XLVII*, 2001, Cracow
- [5] C. Grigoras, “Forensic Audio System”, *IAFP Annual Conference*, 2002, Moscow
- [6] C. Grigoras, “Digital Audio Evidence Analysis”, *Advanced Topics in Electrical Engineering Symposium*, 2002, Bucharest
- [7] UCPTE, *Summary of the current operating principles of the UCPTE*, Text approved by the Steering Committee on 28<sup>th</sup> October 1998