Proceedings of the
35th International Conference on Information Technologies (InfoTech-2021)
IEEE Conference, Rec. # 52438, 16-17 September 2021, Bulgaria

AN ELECTRIC NETWORK FREQUENCY ANALYSIS TECHNOLOGY DEMONSTRATOR FOR EDUCATIONAL PURPOSES

Drл Karl O. Jones, Lewis Hamilton, Dr. David L. Ellis, Colin Robinson, Jago T. Reed-Jones and Kay Morrison.

School of Engineering, Liverpool John Moores University, Liverpool,
e-mails: k.o.jones@ljmu.ac.uk
United Kingdom

Abstract: Authenticating digital audio is a crucial task for audio forensic technicians (AFT) owing to increased use of digital multimedia in litigation. Analysing the electric network frequency (ENF), which can be unintentionally embedded when recording digital audio, is a critical authentication tool. Increasing use of digital media in court means growing demand for AFTs and raised demand for education. Thus, ENF analysis should be key in the education of future audio forensic technicians. A device for educational purposes has been designed to demonstrate the technology and procedures involved in ENF analysis, providing future AFTs practical experience in a key authentication technique.
Key words: Audio forensics, electronic network frequency.

1. INTRODUCTION

In the so-called digital age, the use of digital multimedia as evidence in civil and criminal court cases is an inevitability. With an increased reliance on technology in day-to-day life, it would not be unlikely for the use of cyber, video, and audio evidence in criminal and civil courts to follow an upward trend in coming years. The authentication of digital evidence is key for maintaining its credibility in court, particularly when considering the widespread availability of video and audio editing software. There are a number of methods an audio forensic examiner can utilise for authenticating digital audio evidence, one of which is the analysis of the Electric Network Frequency (ENF) [1]. This innovative method relies on the natural embedding of a low frequency signal within a digital audio file induced by recording the audio within close proximity to cables/devices carrying mains electricity, or by powering an audio recording device from a mains source with poor supply regulation, amongst other factors. This signal is sometimes referred to as ‘hum’ by audio professionals.

In the United Kingdom and mainland Europe, mains electricity has a fundamental frequency of 50 Hz. However, this frequency is not static; it deviates slightly over time resulting from the ever-changing load across the National Grid in relation to the alternating supply and demand over time. Generators situated within a grid network operate in synchronisation, meaning the ENF signal will theoretically be the same across the grid regardless of geographical location, as long as the total of load and loss across the network is equal to generation [2]. By recording and storing these fluctuations as a tabular dataset alongside accurate time and date information, the possibility of analysing the collected mains data against the hum signal embedded within a suspect audio recording arises. If there is a direct correlation to a recorded time-period, the time and date the audio was created can be determined. Furthermore, the hum can be used to determine if an audio recording has been edited by detecting any abnormal variations in the correlation. In the UK, ENF data is published monthly by the National Grid Company (NGC). Each dataset contains mains frequency information collected every second of every day of every month [3], recorded to 3 decimal places.

Owing to the construction of different electric grid networks, it is also possible to localise the creation of an audio recording in specific cases. Unlike in the UK, the United States has three different electric networks for the East Coast, West Coast, and Texas, which operate at a fundamental frequency of 60 Hz. Access to three separate grid datasets within the same nation gives an audio analyst the power to not only determine the time and date of when a piece of audio was created, but to also narrow down the geographical location the audio was created to the three separate geographical zones [4].

Generally, the field of audio forensics is largely inconspicuous despite its ever-increasing importance, particularly in the education sector. There is currently only one university course in the UK (and possibly Europe) that specialises in teaching audio forensics. It is therefore vital that, to meet future demand for audio forensic experts, the system of educating and training future specialists at degree level is robust and effective. Opportunities for students to learn the concepts of ENF analysis are currently restricted to theoretical practices, unlike other audio forensic techniques, where learning can be supported through the practical use of various software packages.

This work aims to create the means for education institutions to implement a low-cost and effective system to allow students to engage with the practical concepts of the ENF analysis technique. The system consists of a micro-computer and an effective low-noise electrical circuit, which will collect mains frequency data with corresponding time and date information, and automatically save the collected material in a cloud-based storage facility. In effect, the device can be classified as a technology demonstrator for use in an educational environment.

2. KEY PRINCIPLES OF THE ENF CRITERION

There have been several publications based on ENF analysis, each offering their own view on the most effective and most efficient methods for implementing the technique within a forensic audio laboratory. Grigoras [1] is considered to be the seminal work, with his findings supporting the concept of mains frequency data being equal across a single network regardless of size. By collecting mains data in three Romanian cities connected to the same grid with a distance covering over 200km, the results were found to be identical across the network.

The work also revealed the extent of the deviations in the grid frequency. According to Grigoras [1], the frequency can deviate from 50Hz by ±0.6Hz for up to 10 seconds. This is useful when considering the rate required for sampling the mains information. The Nyquist sampling theorem states that the sampling frequency must be at least twice the highest frequency within the sampled signal to be accurately represented [5]. Therefore, to represent a signal with a maximum frequency of 50.6Hz, a minimum sample rate of 101.2Hz would be required.

Grigoras [1] also presents useful information in terms of the extraction of the ENF information from a suspect audio file, and the subsequent comparison to the database of mains frequencies. This requires a powerful audio editing software package with Fast Fourier Transform (FFT) capabilities. The evidential recording is first down sampled to 120 Hz, effectively compressing the audio signal by reducing the sample rate and bit depth. Ultimately, down sampling reduces the quality of digital audio, however, this does not affect the forensic process; the purpose for down sampling is likely to be to eliminate noise, as well as to conserve computer memory and increase processing speed.

After reducing the sample rate, Grigoras [1] suggests creating a band-pass filter within the editing software equalisation (EQ) module. A band-pass filter will allow frequencies within a specified ‘band’ to pass through the module, whilst attenuating all other frequencies outside it. In this case, Grigoras [1] specified the band to include frequencies in the range of 49-51Hz with a slope of 24 dB/octave. This isolates the ENF signal. The processed signal was then plotted as a spectrogram with an FFT of 4096 points and visually compared to the ENF database. Cooper [6] describes the physical process of collecting mains data, involving using a transformer to step-down the 240V rms signal to a secondary working voltage. The signal is then directed through a pulse generator, converting the signal to a square wave before being recorded on a PC soundcard.

There are some potential flaws in the ENF approach, for example the criterion proposed by Grigoras [1] has a reliability issue at the extraction and matching stage. The ENF signal had a slight offset when compared to the database. This was particularly prevalent in recordings with a length of 10 minutes or less. A suggestion was made to implement a maximum correlation coefficient matching procedure, based on analysing the equivalence of shape plotted as a spectrogram [7]. Having established that obtaining a precise match can depend on the length of recordings, it is important to examine the limitations audio signal poses. Research suggests that the recording device and recording environment affect the prominence of ENF signals embedded in an audio file [8].

Environmental factors can include wave interference and movement of the recording device. Hajj-Amad *et al*. [8] carried out experiments suggesting that if more than one sound source is emitted at the time of recording, there is potential for destructive interference from compressions and refractions having a cancelling effect. This can cause the ENF signal to appear at different strengths or not appear at all. Movement of the device may cause ENF traces to become distorted and indistinguishable. It is also suggested that the signal-to-noise ratio (SNR) of the recording device can affect the strength of the ENF signal. If a device has a particularly low SNR, the ENF signal (which would typically be classified as noise) is likely to be more prominent, and vice versa.

3. ENF CAPTURE DEVICE DESIGN

3.1. Raspberry Pi

For the frequency of the mains to be accurately determined, the mains signal must be registered by a form of computer device with suitable software development. A Raspberry Pi (RPi) was chosen because it is arguably the most popular single-board computer on the market. Originally released in 2012, RPi is arguably the most popular single-board computer on the market. A key RPi feature is the 40 General-Purpose Input/Output (GPIO) pins, where each pin can be designated in software as either input or output. These allow for a maximum 3.3 V signal to be registered in or out of the RPi, creating a channel for the stepped-down mains signal.

The ability to record time and date information in real-time is crucial in the context of collecting periodic data for use in forensic investigations. The RPi 4 does not have a built-in real-time clock, thus the RPi was connected an external real-time clock (RTC) chip to the spare GPIO pins. A variation of the DS1307 chip, designed specifically for the RPi, provided an accurate method for receiving time and date.

3.2. Circuit Design

A key consideration is the method to capture the signal from the mains for the precise frequency to be calculated. To calculate the ENF, the signal from the mains will need to be fed to a DC-supplied computer device, thus a method for converting the alternating mains current to DC needs to be devised, such as through rectification.

The circuit designed was based on the HCPL-3700 AC/DC to Logic Interface optocoupler, since the optocoupler removes the possibility of noise, voltage surges, and ground loops. It is also effective for removing of any electric or electromagnetic interference, a crucial element for capturing the mains signal. The HCPL-3700 chip features a diode bridge to obtain compatibility for use with AC signals. Internal hysteresis is also included within the chip for extra noise and switching immunity.

The LED used to transmit information is protected from over-current and voltage issues with internal clamping diodes, contributing greatly to the reliability and functionality of the chip for ENF capture.

The design incorporates a transformer to act as the input method for the AC mains signal and to step-down the signal to a suitable working voltage (Fig. 1). Despite the typical use of an optocoupler device to provide isolation, in this case the transformer carries out this role. The HCPL-3700 is instead incorporated as a convenient method for extracting zero crossings, in the form of the logic output derived from the switching across the threshold.

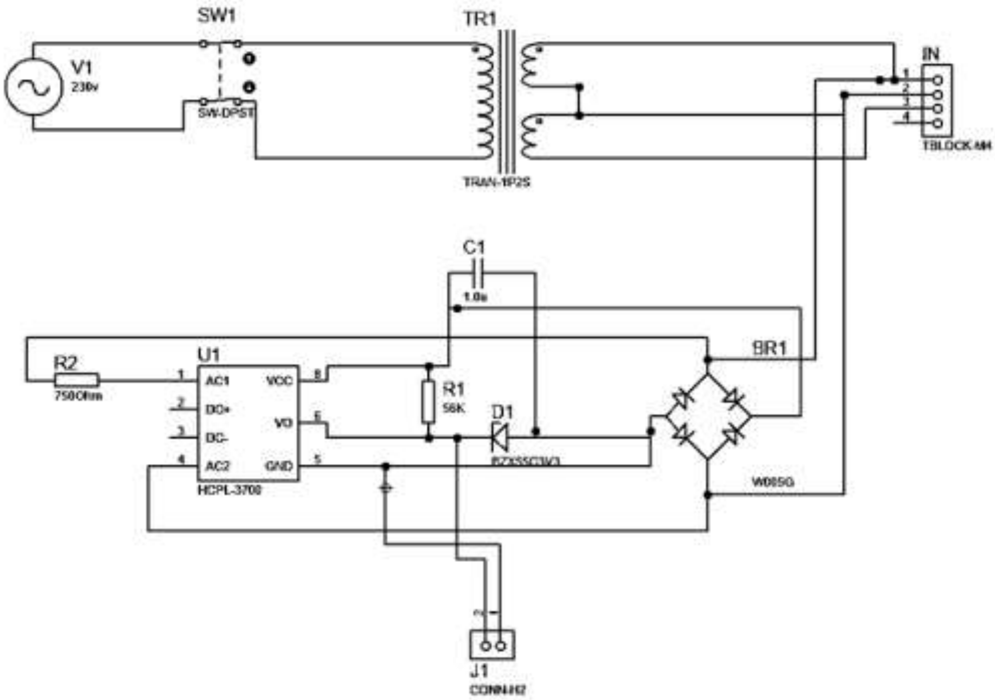

Fig. 1. Transformer and Optocoupler circuit

The optocoupler device was designed within software and simulated accordingly. A square wave signal was shown to be alternating between logic 1 and 0 with a peak of 2.8 V. Each cycle was produced with sharp straight edges, creating a signal ideal for use with edge detection-based software. Analysis of the output signal showed some slight spikes in volage at each cycle at the point of logic 1; however, this was judged to be within an acceptable range since there was no clear effect on the purity of the rising/falling edges.

3.3. Software

Python was the chosen language owing to its compatibility with RPi, and its capability of implementing both the zero-crossing detection and FFT methods of frequency estimation. The capability of counting rising/falling edges on the GPIO pins of the RPi is an interesting prospect. This requires little processing power and can be extremely accurate if implemented correctly. The simplicity of this method also holds an advantage over calculating the FFT of the signal.

The signal registered by the GPIO must then be processed for edge detection to be carried out and counted, which will in turn enable a frequency estimation to be determined via a reliable algorithm. The final requirement of the software is to record each estimation in a file format alongside time and date information, in one second intervals. The fie is then be uploaded to the cloud once data is halted and the file closes.

An exemplar of the Python functions code is provided in Fig. *2*. When a rising edge is detected, the function '*crossings*' increments the counter by a value of 1 for each edge detected, whilst simultaneously saving a timestamp for each edge in the 'stamps' list using the *.append* method [9].

```
def crossings(channel):
    global count

    count+= 1
    stamps.append(time.time())

def status():
    global count

    d_t = time.strftime("%d/%m/%Y %H:%M:%S")
    frq = round(1/mean(diff(stamps)),3)

    import csv
    with open("Oct2020.csv", "a") as csvfile:
        datawriter=csv.writer(csvfile)

        datawriter.writerow(["Date:", d_t, "F:", frq, "Hz"])


    print(d_t, frq)
    count=0
    stamps.clear()
```

Fig. 2. Functions for calculating frequency from zero crossings.

The function, '*status*', calculates frequency in hertz and writes the result to a CSV file. Time and date information is also written alongside the frequency value. The frequency estimation algorithm calculates the value of T as an average cycle period based on the differences in time between every detected rising edge. This eliminates the requirement to make fast calculations, and thus decreases the likelihood of timing errors occurring. The print command in the 'status' function will display the information which is saved in the CSV file on screen as a reference. The counter is then reset to zero and the values in the timestamp list cleared to be restarted for the estimation of the next seconds' frequency.

4. PRACTICAL IMPLEMENTATION

4.1. Recording Mains Frequency Method

To demonstrate the applicability of the detection system in carrying out the forensic process of ENF analysis, the RPi/optocoupler ENF device collected data over an hour-long period from 18:00 on 31st October 2020, so that an attempt could be made to match the ENF data contained within the RPi dataset with data from the National Grid. Additionally, an exemplar audio recording was created during the RPi data collection period, the aim being to confirm precisely at which point was recorded during the hour. In order to ensure that the 'hum' from the electronic mains was captured within the audio file, the recording device was placed in proximity to a pair of ground looped high-fidelity monitors with connection to the mains. This effectively amplifies the 50Hz hum, as well as the associated harmonics, and embed the frequencies within the audio file. The audio file was recorded between 18:23:00 and 18:40:00 [7]. An Apple iPhone X was used as the audio capture device using a compressed .M4A format.

The audio file was converted to a 16-bit 44.1 kHz lossless wave file (*Audio1.wav*), to replicate the duplicate file a forensic examiner would typically make in a real-life scenario to perform the processing on, in order to preserve the evidence in the original recording. It is crucial that the conversion is never to a lossy file format; data loss has the potential to harm the credibility of the resulting evidence in litigation. The processes which followed were largely based on the recommendations by Grigoras and Smith [10] for extracting the ENF.

Although in a real scenario, the forensic examiner would likely have an automated system in place to extract and match the data, it was judged to be beneficial in this scenario for a manual process to take place since it allows students to be fully aware of the entire process. Undertaking the extraction and matching manually may be considered a long and arduous process, however this may be a superior method for assisting the learning of students by developing their understanding of the process to a further extent. *Audio1.wav* was loaded into the Audacity audio editing software [11]. The file was then down sampled to 120Hz [10], which results in the bulk of unwanted noise to be eliminated. The next stage was to apply a high-pass filter with a 24dB roll-off, with a cut-off frequency of 49Hz, and then a 24dB roll-off low-pass filter with a cut-off frequency of 51Hz, to eliminate any residual noise around the fundamental frequency. The file was the exported as a 16-bit WAV PCM.

The NGC data downloaded for the electricity network is recorded in the time domain, which presents a formidable challenge when considering frequency data stored within the audio is represented in the frequency domain. To convert the data from the frequency to the time domain, ideally a short-time Fourier transform (STFT) algorithm should be used [12], using the following equation:

$$X[n,\lambda] = \sum_{m=-\infty}^{\infty} x[n+m]w[m]e^{-j\lambda m} \qquad (1)$$

where $x[n]$ is the signal, $w[n]$ is the window length and λ the frequency.

In order to extract the frequency information from the audio recording, the *Audio1.wav* file was loaded into iZotope RX audio restoration software [13] and a random sample of 1 minute in length was taken. Owing to the short time span of the 1-minute sample, the audio could be split into windows of equal length, and frequency for each window calculated manually via the FFT spectrum analyser module in RX7. The data from each window could then be manually inputted to Microsoft Excel to be plotted. The reasoning behind this method was based on the principles of the STFT. It must be reiterated that this method of ENF extraction is unsustainable and is potentially extremely labour intensive if audio files of significant length are presented; therefore, for a practical commercial forensic implementation the development of a suitable STFT algorithm would be highly beneficial. However, in terms of an educational exercise then this approach is of value for students to appreciate the whole analysis process.

4.2. Results

An initial visual comparison of the RPi data and the NGC data shows a very good match, particularly in the shape of the curve (Fig. 3 and Fig. 4). The actual RPi recorded data did not precisely match the NGC provided data, however the difference between the two values at each sample is small being in the range ±0.03% (Fig. 5). A possible explanation for this would be the rounding function which is present in the RPi frequency estimation algorithm. A Pearson correlation coefficient was calculated to be 0.9171 which indicates a strong degree of correlation. From a forensic perspective, since the primary function is the determine the time period an audio recording was made, having a small error for each sample is not a significant issue since the overall match of actual frequencies compared to recorded frequencies is the critical factor including the contour of the frequency changes (shape of the curve) given that an audio recording will last longer than a couple of seconds/samples.

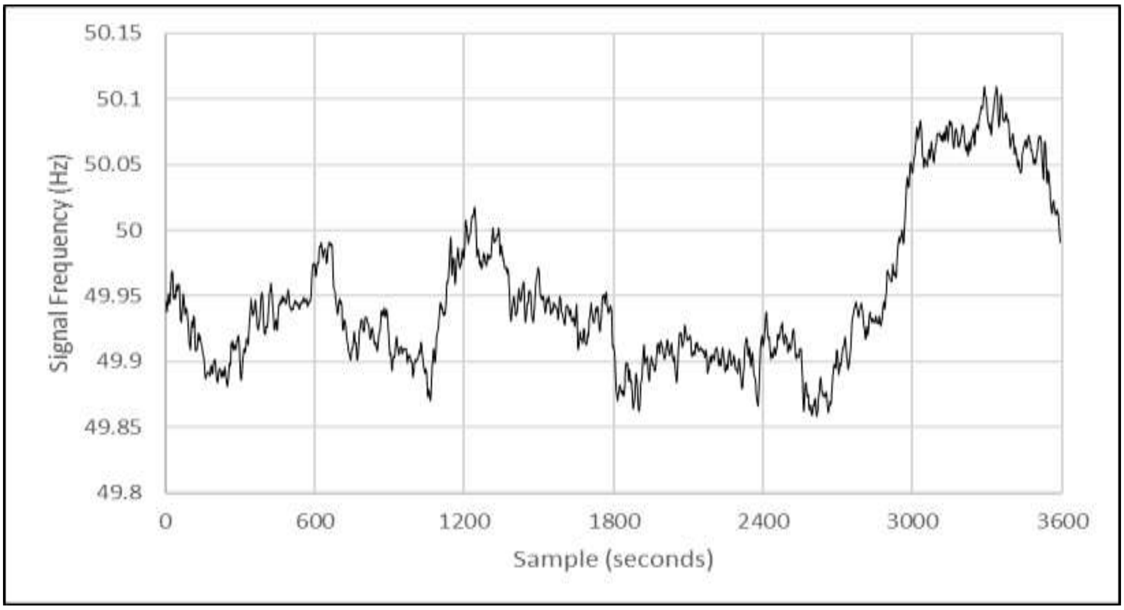

Fig. 3. National Grid Company electrical network frequency data

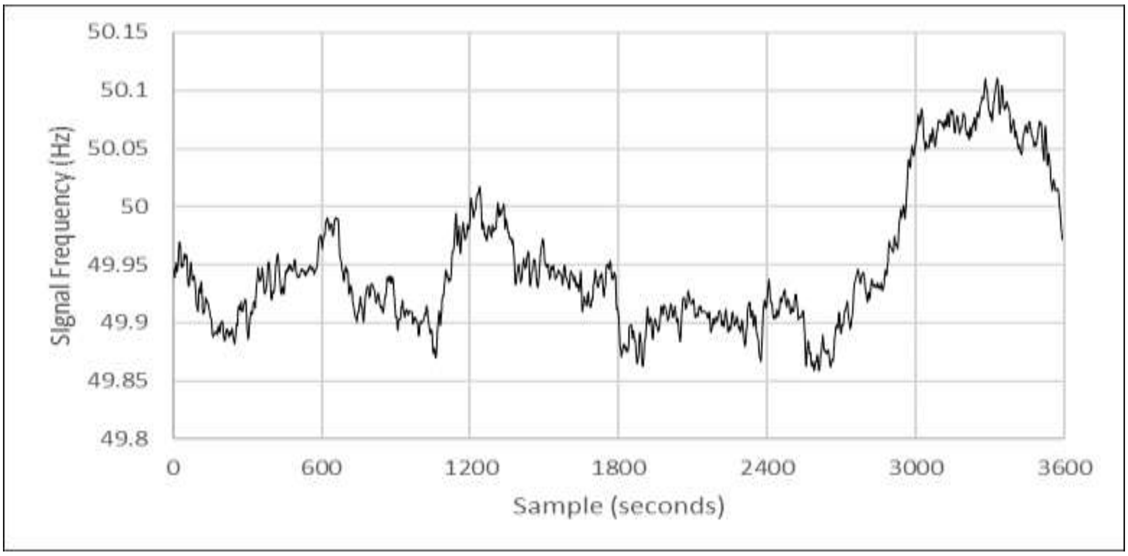

Fig. 4. Raspberry Pi device recorded frequency

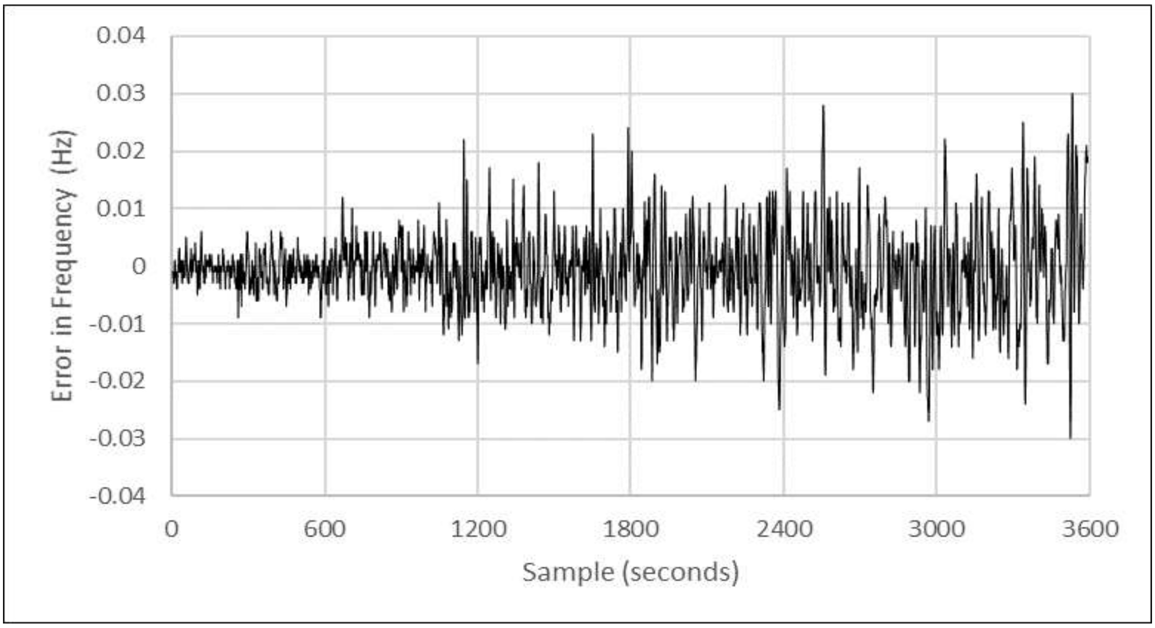

Fig. 5. Difference between NGC and Raspberry Pi frequency values

In relation to the audio recording made on the iPhone, the frequency data points extracted from 1 minute from the 17 minutes of *Audio1.wav* were subsequently plotted and compared against the RPi data and the NGC dataset. The *Audio1.wav* extracted frequency data is shown in Fig. **6**, the RPi data is shown in Fig. 7 and the NGC data is given in Fig. 8. Correlation coefficient of 0.88 was calculated for the *Audio1.wav* frequency values against the RPi determined frequencies, while the coefficient was 0.96 for the correlation between extracted Audio1.wav data and NGC values. Clearly the *Audio1.wav* extracted ENF data had a slightly better correlation with the NGC database than with the RPi determined data.

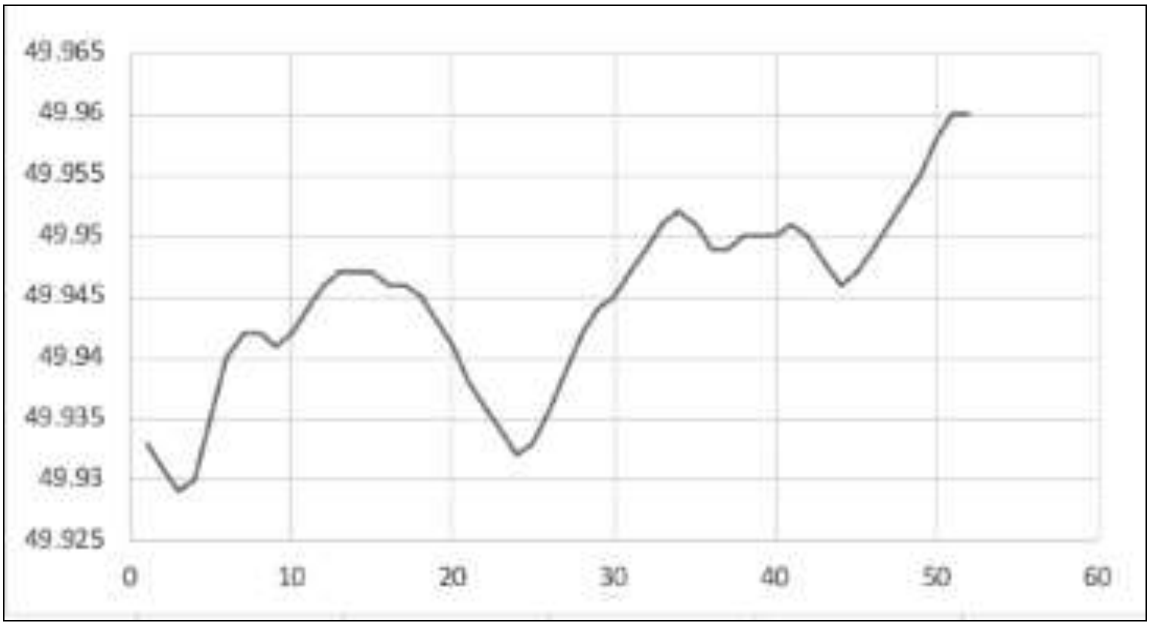

Fig. 6. ENF extracted from 1 minute of the audio file 'Audio1.wav'.

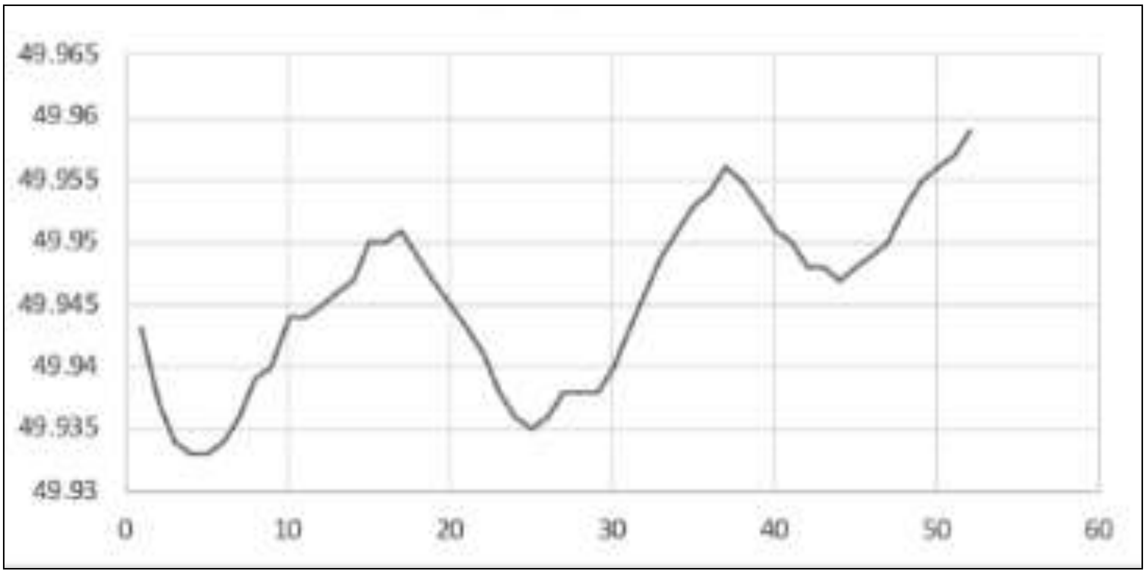

Fig. 7. ENF values obtained from the RPi device.

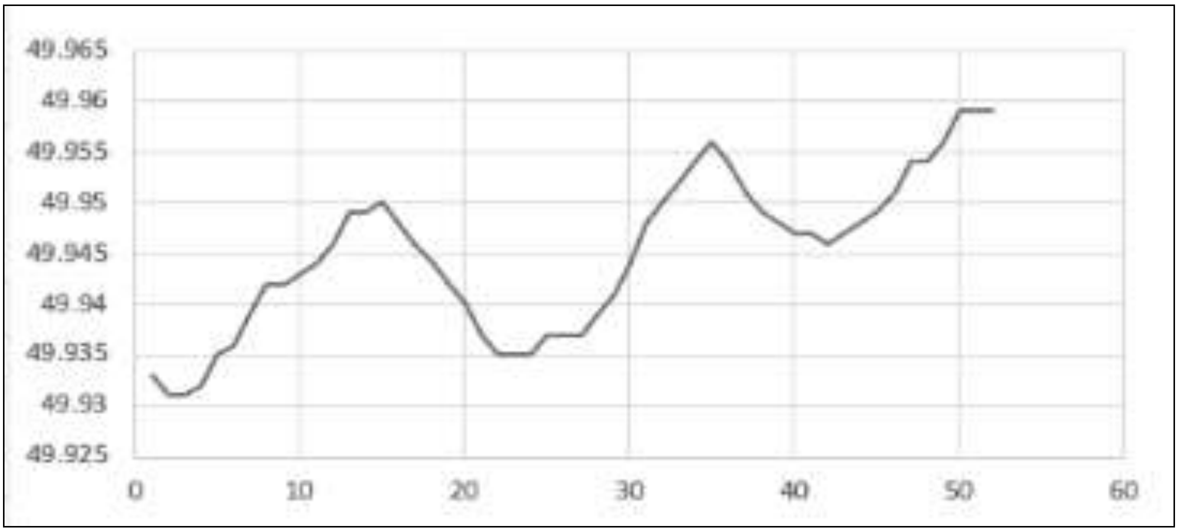

Fig. 8. ENF values downloaded from the National Grid Company database.

5. CONCLUSIONS

The described Raspberry Pi ENF device demonstrates its compatibility with the existing National Grid database by collecting data samples. The RPi device can obtain highly accurate data from the electrical mains, except for some minor errors for individual samples. The overall correlation between the RPi and National Grid data reinforces the argument that the device is highly accurate, with a coefficient of >0.9 [14].

A key aim of the work was to demonstrate how the device could be used in education to demonstrate and practice the processes of extraction of ENF information from an audio recording and matching it to collected ENF data (either from NGC data or from a dedicated device such as the RPi device described here). The process of preparing an audio file for ENF extraction proved to be successful in isolating the fundamental frequency of the 'hum', leading to more accurate frequencies to be extracted. This has been shown through the forensic analysis of the recorded audio file

It can be concluded that the device would be compatible for use in an educational setting since the technology and principles of the data collection aspect of ENF analysis can be demonstrated to a high level. However, in its current state, the device requires further development to render it fully operational for students to use as a practical tool when learning how to authenticate digital audio.

REFERENCES

[1] Grigoras, C., 2003. *Digital Audio Recording Analysis: The Electric Network Frequency Criterion,* Hibernia: Diamond Cut Productions, Inc.

[2] Sidhu, T., 1999. Accurate Measurement of Power System Frequency Using a Digital Signal Processing Technique. *IEEE Transactions on Instrumentation and Measurement*, 48(1), pp. 75-81.

[3] National Grid ESO, 2019. *System Frequency*. data.nationalgrideso.com/system/system-frequency-data

[4] Zjalic, J., Grigoras, C. & Smith, J. M., 2017. *A Low Cost, Cloud Based, Portable, Remote ENF System.* Arlington, Audio Engineering Society.

[5] Oshana, R., 2006. Overview of Digital Signal Processing Algorithms. In: *DSP Software Development Techniques for Embedded and Real-Time Systems.* Burlington: Newnes, p. 66.

[6] Cooper, A. J., 2008. *The Electric Network Frequency (ENF) as an Aid to Authenticating Forensic Digital Audio Recordings - An Automated Approach.*, Denver: Audio Engineering Society.

[7] Geradts, Z. & Huijbregste, M., 2009. *Using the ENF criterion for determining the time of recording of short digital audio recordings* , Den Haag: Netherlands Forensic Institute.

[8] Hajj-Amad, A. et al., 2018. *Factors Affecting ENF Capture in Audio,* Maryland: IEEE.

[9] Singh, A. & Parida, S., 2019. *Power System Frequency and Phase Angle Measurement using Raspberry Pi,* Patna: Indian Institute of Technology.

[10] Grigoras, C. & Smith, J. M., 2012. *Advances in ENF Analysis for Digital Media Authentication.* Denver, Audio Engineering Society.

[11] https://www.audacityteam.org/

[12] Selesnick, I. W., 2009. *Short-Time Fourier Transform and Its Inverse. New York University.* eeweb.engineering.nyu.edu/iselesni/EL713/STFT/stft_inverse.pdf

[13] https://www.izotope.com/en/shop/rx-8-standard.html

[14] Jaadi, Z., 2019. *Everything You Need to Know About Interpreting Correlations.* towardsdatascience.com/eveything-you-need-to-know-about-interpreting-correlations-2c485841c0b8