IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 18 November 2022, accepted 27 December 2022, date of publication 4 January 2023, date of current version 9 January 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3234228

RESEARCH ARTICLE

An Ultra-Low Energy Solution for Large-Scale Distributed Audio Monitoring

STEVEN FENTON
School of Computing and Engineering, University of Huddersfield, Queensgate Campus, HD1 3DH Huddersfield, U.K.
e-mail: s.m.fenton@hud.ac.uk

ABSTRACT This paper presents a flexible, low-cost, and ultra-low-energy acoustic sensor network solution. This enables both highly distributed and real-time audio measurements to be employed in a wide variety of scenarios. These types of distributed measurements can be useful in environmental monitoring, healthcare settings, smart cities, and security surveillance systems. However, they are typically inflexible in their applications and require specialized deployment. The proposed solution is nonintrusive and uses long-range radio to enable low-power distributed audio monitoring. The system can recognize and be configured for different types of audio measurement and analysis tasks (such as sound pressure level, annoyance, direction of arrival, and roughness), and is shown to be as accurate $r=0.998$ as a high-cost measurement system. Remote control, time of day, and grid capture mechanisms can also be realized. The proposed solution provides the ability for an over-the-air configuration, leading to minimal physical access required once deployed. In addition, higher data rates are possible, offering real-time monitoring metrics.

INDEX TERMS Acoustic sensor networks, edge computing, energy efficiency, remote health monitoring, smart cities.

I. INTRODUCTION

A. MOTIVATION FOR THE WORK

Long-term exposure to unwanted noise (also referred to as environmental noise) can cause various negative health effects. These include annoyance, depression and anxiety [1], sleep disturbance [2], cardiovascular issues [3] and cognitive impairment [4]. Based on these observations, it's with an urgency that solutions are developed that are both affordable, robust, energy-efficient, and configurable to meet the demands of a distributed and large-scale audio monitoring system.

Such a system would allow the collection and analysis of audio data, which could be utilized to control and prevent excessive noise exposure. This would result in human benefits across many scenarios, such as healthcare settings, urban environments, and schools. Affordability in low-power sensors and recent developments in system-on-chip (SOC) solutions make the possibility of edge computing a reality and can significantly reduce the data traffic associated with audio-based systems.

In ever-crowded environments, such as cities, classrooms, and hospitals, proximity to sound sources is inevitable; therefore, exposure to higher intensity is increasingly likely. The UN has predicted that 68% of the world's population will live in urban areas by 2050. This represents an upswing of 54% since 2016 [5].

The structure of sound, including its level, has a bearing on how annoying it is perceived by the listener [6]. For example, the scraping of fingernails on a whiteboard is perceived to be significantly different from the sound of a babbling stream. Some sounds can be considered as unwanted noise if no discernible information is present in the auditory stream.

Unwanted noise sources are considered more annoying than discernible sources when auditioned at the same intensity. Exposure to discernible sound sources at high levels of intensity is likely to cause the source to be considered unwanted noise unless a conscious choice is made towards its exposure by the listener [7].

Sound sources are all around us, differing in both their harmonic structure and the level at which they are exposed.

The associate editor coordinating the review of this manuscript and approving it for publication was Shuangqing Wei.

Taking both of these in isolation, the harmonic structure can be thought of as the meaning of sound, for example, a bird song or a pneumatic drill. The sound level of the source to which we are exposed is often associated with the perception of its size or proximity to that source [8]. If a sound source is recognized (remembered) as small (e.g., a bird), high-intensity exposure would indicate closeness to that source, rather than the presence of a giant bird.

B. CHALLENGES

The assessment of exposure to sound sources, whether desired or unwanted, presents a unique set of challenges. For example, as outlined previously, some sounds are more annoying than others, and the time of day at which exposure occurs has an impact on what is deemed acceptable. Legislative policies have been suggested by the European Parliament [10] to help assess and manage environmental noise across urbanized areas in member states.

The document qualifies noise issues as local; therefore, the management to minimize exposure to sound levels is down to the individual member state. The EC directive also stipulates that noise-level indicators should be utilized in noise planning. These are the Lden (Day-Evening-Night level in dB) to assess annoyance and Lnight to assess sleep disturbances. Noise exposure limits are determined locally.

It is worth stating that the measures outlined are based on A-weighted average sound levels defined in ISO 1996-2:1987 [11], which is now superseded by ISO 1996-2:2017 [12]. Where member states have no national computational methods or wish to change a method, additional standards are recommended [10]. Therefore, it is imperative to propose a solution that allows local configurations for measurement standards and methodologies in addition to the acceptance parameters.

Ease of deployment is imperative, as the collection of measured data must also be performed within grid sizes and heights and with consideration of reflective surfaces, as defined in the ISO standard [12]. This can make the actual process of collection, calibration, and correction time consuming and expensive. A study by Escobar et al. [13], which examined the grid sampling method, required approximately two months to map an area of $10km^2$. In their study, multiple measurements were performed between an 8 am-8 pm interval. This approach makes it impossible to capture simultaneous measurements across a measurement grid. Furthermore, if the measurements need to be repeated, the manual grid sampling must be repeated.

The equipment used in the collection of noise data is typically made up of either a portable recording system, which requires a lot of manual intervention, or fixed wireless-based solutions [14]. The latter example utilizes WiFi, which, while offering a means of capturing automated measurements, is not the most efficient in terms of the transmission distance to the access point versus power.

Current portable measurement solutions are generally not ideal for continuous monitoring because of their size and cost, which is the case in both external and internal measurement studies. Typical internal studies of noise exposure include monitoring classrooms [15] and hospital noise levels [16]. Noise in these cases (and in other studies) has been shown to affect cognitive performance and patient recovery. Roberts et al. [17] suggested a monitoring strategy for better patient care by integrating noise monitoring solutions into building management systems. Similar to city mapping, they suggested sound mapping of zones throughout the hospital. They also specified the key requirements to consider when selecting a noise monitoring solution. The successful implementation of such solutions requires a portable, low-cost, and nonintrusive approach.

The authors of [18] described an audio sensor network employed in a telemedicine application that, although able to perform local event detection, comes at the expense of its power requirements and bulk. This outlines another advantage of the system proposed in this paper, that is, the use of microelectromechanical systems (MEMS) [19] for acoustic signal capture. Advancements in transducer technology have enabled extremely low-power, compact, and accurate audio measurements to take place. In this paper, the performance of MEMS was shown to match that of more expensive and bulkier transducer technology.

Data size requirements can be considerable when dealing with audio samples, particularly when accurate high-sample-rate capture requirements are necessary. The transmission of such data in real time can be very energy-intensive; therefore, low-power edge-based solutions are preferable because they can perform smart local processing to drastically reduce the data payload [20]. Smart data can be used to augment existing systems (such as videos) to improve overall system performance [21]. For example, the proposed system shows how a loudness measure, along with other node parameters, can be transmitted using only four bytes of data.

The authors of [22] demonstrated that it is possible to achieve real-time voice over an acoustic sensor network. Their work utilized sensors capable of 50m-100m transmission, while using multi-hop mesh transmission. The system proposed in this paper incorporates the use of Long-Range Radio (LoRA). The distances LoRa signals can travel exceed 700km. Generally however, LoRa communications range from up to three miles (5 km) in urban areas, and up to 10 miles (15km) in more rural areas [9]. By incorporating a multi-hop mesh transmission into the proposed system architecture, vast areas of measurement can be achieved. Acoustic sensors also can be used for classification and recognition [23], [24]. Audio classification-based systems often lean towards neural networks, and despite the conflicting memory requirements of this approach with that of typical embedded system topologies, significant efforts are being made to overcome memory hurdles. In [25], a system for distributed audio classification using sparse representation over the cloud for IoT was proposed. In that application, audio classification is done separately on different nodes using distributed sparse representation. This avoids any centralized processing and a reduction in the data throughput. In [26], EdgeL3, a 95% sparsified version of L3-Net, was proposed as the first reference model to introduce state-of-the-art robust machine listening to the edge. In essence, they significantly reduced the parameters utilized by the CNN to reduce its overall memory footprint with a minimal reduction in performance accuracy. The aim of this paper is to propose a low-power solution that can leverage these types of classifications, and audio event topologies providing sufficient local memory are available to facilitate inference.

C. CONTRIBUTIONS

In this paper, a flexible, low-cost, ultra-low-energy solution that would enable both highly distributed and real-time audio measurements to take place is presented. The proposed solution is nonintrusive and utilizes LoRa [9] to enable long-range and distributed audio monitoring. The system can recognize and be configured for different types of audio measurements and edge-based analyses (such as sound pressure level, annoyance, direction of arrival, and roughness). Remote control, time-of-day, and grid capture mechanisms can also be used. Furthermore, the proposed solution can switch to Wi-Fi, which can enable an over-the-air configuration; therefore, minimal physical access is required. Higher data rates in this mode enable the real-time monitoring of metrics. The solution can be configured via its back-end to be predictive and therefore, preventative, thereby providing an effective noise management system.

It's demonstrated that,

- The proposed wireless sensor network can achieve acoustic equivalence with current solutions at a much lower cost.
- The system has a very small form factor compared to current solutions and offers significant savings in energy usage.
- The system is configurable in terms of both application and edge-based processing methodologies, thus allowing local configuration of the measurement standards and acceptance parameters.
- The system demonstrates significant savings in the data payload, thus reducing both transmission time and energy usage.

In Section II, the architecture of the proposed system is described along with details of the prototype measurement node and edge processing capability. Section III outlines the testing and underlying performance of the prototype and presents an example of a prototype remote monitoring interface. Section IV concludes the paper and outlines avenues for future work.

II. SYSTEM PROPOSAL

As shown in Fig. 1, the proposed system comprises a set of distributed monitoring nodes. These sensor nodes, also known as motes, can gather, process, and communicate with other connected nodes in the network. They're also capable of reconfiguration and adapting to each application.

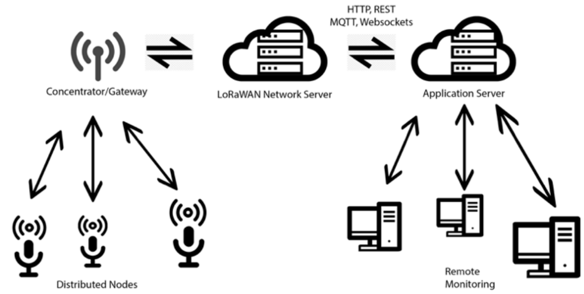

FIGURE 1. System Proposal.

Each monitoring node can utilize a variety of wireless communication methods depending on the required mode of operation. These methods include LoRa[9] and WiFi[27]. Bluetooth[28], each of which is available and can be used to enable different communication topologies and strategies as required. In the test system, mesh communication is not utilized. Instead, each node communicated to a central and remote server via a gateway at periodic intervals, as denoted by the application.

Data can be accessed remotely for analysis on the client system, in this test case, this was achieved through the use of an off-the-shelf MQTT-based client application called TagIO [29]. The client system can incorporate real-time or batch-based processing, which enables AI-based models to predict future audio events or to propose optimization schemes to minimize exposure to certain audio attributes.

LoRa was chosen as the default communication method for each measurement node because of the requirement of providing a low-energy solution. This allows the greatest range between nodes, while offering extremely low power consumption. In addition, the LoRaWAN protocol allows easy integration with the Internet, allowing a network of nodes to be built quickly and efficiently across the existing architectures.

To facilitate network integration, the LoRa Gateway enabled each node to communicate with the cloud network. Testing was performed using the Things Network (TNN) [30] to facilitate data packet collection, and TagoIO was used to build an application interface to enable the remote monitoring of measurement data. The proposed method is not limited to this network and its applications, as other commercially and equally viable solutions are available. TNN and TagIO provided a completely free platform for the proof of concept.

A. MEASUREMENT NODES

The nodes comprise a low-energy, high-performance microcontroller unit (MCU) that can perform edge-based functionality in addition to audio capture, battery monitoring, and other sensor-based tasks. Edge-based processing reduces data requirements when transmitting to the server, thereby lowering the overall energy requirements of the system. At the point where the LoRa nodes within the system become activated and transmit, the maximum energy is utilized, and represents the point at which the maximum battery drain

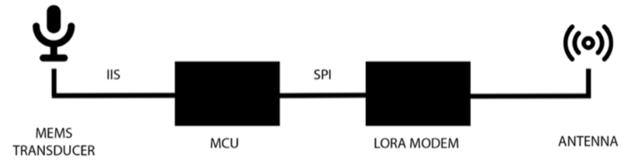

FIGURE 2. Measurement Node.

occurs. Therefore, a reduction in payload leads to a reduction in energy and an increase in battery life. These energy savings are demonstrated later in this paper. Additionally, nodes can be reconfigured remotely to adapt to various processing requirements, thereby offering a fluid solution after deployment. For example, where measurement standards differ across regions, or the application of node processing is modified from loudness measurement to audio event classification. Additional processing such as correction based on reflective surfaces [12] could also be employed as required.

The acoustic transducer used in the proposed node is a MEMS microphone [31]. The use of a MEMS type offers ease of implementation via the IIS protocol [32] and direct support by the MCU with minimal external components. In addition, the transducer is very robust against water and dust ingress and can operate with very little current. The prototype node did not employ any additional protection against rain, snow, or humidity changes. However, the transducer was mounted on a small PCB following manufacturer guidelines to ensure maximum protection and a bottom-port vent hole.

The transducer used was the Invensense INMP44, which has an extremely compact form-factor measuring $4.72 \times 3.76 \times 1$ mm in a surface-mount package. It is a high-performance, low-power, digital-output, omnidirectional microphone. The device has in-built signal conditioning, an analog-to-digital converter, anti-aliasing, filters, power management, and 24-bit IIS interface. IIS interface allows the INMP441 to connect directly to the MCU without the need for an additional audio codec [30]. The transducer has a high sensitivity at -26dBFS and a wide frequency response which is nominally flat between 60Hz and 15kHz.

The MCU used in the test system was the ESP32 [33]. This complete SOC solution provides Wi-Fi and Bluetooth functionalities with very few external components. LoRa communication was achieved through the interface of LoRa MODEM and SPI using a Semtech SX1276 device [34]. The proposed node is illustrated in Fig. 2.

A collection of nodes can be arbitrarily positioned throughout the area of interest, or in the form of a measurement grid. The grid density would be determined by the application requirements. As concluded by Escobar et al. [13], there can be some variation in city wide measurements based on the grid density chosen, however, the proposed system makes it possible to capture simultaneous measurements across the grid, leading to more accurate monitoring. Furthermore, if the measurements need to be repeated, automatic grid sampling can be performed with no user intervention. Location data within an area can be obtained either via programmed grid coordinates, mapping IDs, or using a Global Positioning System (GPS). The nodes can be self-aware of their locations, thus allowing different configurations to be automatically deployed.

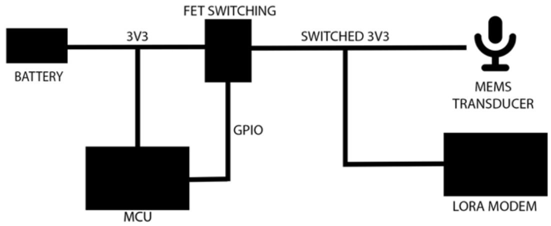

FIGURE 3. Low Power Switching to external peripherals.

The nodes operate fully remotely, storing local measurements until either a request is made for data, or a periodic sending interval is reached. When not monitoring or processing a set of measurements, the nodes enter deep-sleep mode, allowing ultralow current operation. MCU real-time clocks are utilized to wake nodes to measure, send, or process data.

A prototype node was connected to a 1000mAh battery for testing purposes to establish the lifespan of a small-capacity battery. The battery was chosen based on its size, allowing for a very compact, nonintrusive, and fully portable solution to be realized in the test system. The MCU used a Field Effect Transistor (FET) switch connected to one of its General Purpose Input Output (GPIO) lines to allow the switching of power to the external LoRa MODEM and MEMs microphone. This switching mechanism allows the MCU to prevent current from being drawn from the battery by the attached peripherals during its deep-sleep modes. The switched power connections are shown in Fig. 3.

The prototype node incorporated the use of a single MEMs transducer. However, owing to the additional IIS peripheral available on the MCU, four MEMS transducers can be incorporated without additional hardware requirements. This enables features, such as the Direction of Arrival (DOA) and specific listening lobes, to be processed in the sound field. With further integration of additional MEMS transducers, a full 3D sound field analysis could be employed.

This additional 'spatial' based processing has benefits over traditional single microphone capture. The system can detect differences in each audio stream, thus allowing complex spatial information to be extracted and processed, which could also include full 3D triangulation of the sound source, identification, and spatial annoyance factors.

B. EDGE PROCESSING

The MCU chosen in the test systems was a dual-core variant offering high performance and low cost, while having the capability of GPIO switching, deep sleep modes, and floating-point operation. Being an SOC, it greatly reduces the component count of the processing nodes. It is widely available in different variants; thus, system customization is possible.

For prototype purposes, the MCU was programmed around a FreeRTOS kernel [35]. This kernel is freely distributed under the MIT open-source license, making it accessible to all developers. The kernel offers the ability to easily configure the MCU for scheduling and low-power switching. The prototype also utilized the IBM LMIC Library [36] to facilitate the LoRaWAN implementation.

Edge processing was applied by the MCU in the form of an integrated loudness measure Leq(Z), performed across 125mS at half-hourly intervals using (1). This loudness measure takes place across the 125mS interval with no frequency weightings being applied. Additional weighting can be applied using weighting filters [37]. Typically, A-weighted filters are employed to approximate to human hearing sensitivity.

$$Leq\,(W) = 10log_{10}\left[\frac{1}{T}\int_0^T \frac{x_w^2}{x_{Ref}^2}dt\right] dB \tag{1}$$

where:

W represents the frequency weighting, which in this case was unweighted. x_W is the signal at the output of the weighting filter. x_{Ref} is the reference level and T is the length of the audio sequence.

The measurement was performed at 48kHz in 24-bit signed-integer format. The data format was compatible with the transducer IIS output. In addition to the periodic loudness measure, the node was programmed to read its battery level, which, along with two additional bytes of random data, were sent to the collection network via a LoRa gateway concentrator. Random bytes were named annoyance and exposure levels to indicate their possible usage.

Annoyance is somewhat subjective, and the context of the application deployment would determine what constituted an annoyance level depending on factors such as proximity to differing types of sound sources, time, and duration of exposure. The European Parliament [10] specify common noise indicators to assess annoyance The directive also outlines that member states can use supplementary indicators to monitor or control special noise situations. The proposed system can accommodate this.

The equation shown in (2), can be used to implement a basic annoyance indicator. This would allow the display of annoyance above and below the threshold chosen.

$$Annoyance\,(A) = Leq\,(A) - ThreshdB \tag{2}$$

where:

$Leq\,(A)$ represents the frequency weighted integrated loudness level in dB. *Thresh* is the defined annoyance threshold in dB.

Edge processing results in a significantly reduced data payload packet size, consisting of only four bytes. The loudness measurement was stored as a single byte, representing the Leq(Z) level in decibels. This represents a data-byte saving of 18000:1 compared with the transfer of raw audio samples at the same sample rate. The payload is transmitted to the LoRa concentrator, along with the node ID, location data and timestamp in a JSON packet. A smaller packet size decreased the transmission time and significantly reduced the power required during this process. An alternative would be to send raw audio data to the concentrator, which would be highly inefficient, and, in the case of LoRa, would probably violate the network's usage policy.

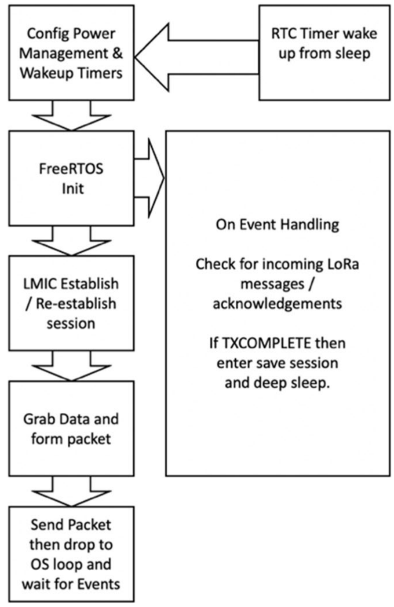

FIGURE 4. Prototype code structure.

If real-time (< half-hourly interval) measurements are required, reconfiguration can be easily achieved at the expense of higher power consumption. If this type of operation is required, the unit would likely require a non-battery-type power supply and utilization of the WiFi capability of the MCU rather than the attached LoRa Modem.

The MCU contains IIS peripherals, which enable seamless connection to the MEMS microphone. In the prototype system, audio samples were transferred from the IIS interface to internal memory using MCU-dedicated DMA controllers. This allows streaming of sampled data without requiring the main CPU of the MCU to copy each data sample. The coding structure is illustrated in Fig. 4.

Upon booting, the MCU sets up power management and wake-up timers. The FreeRTOS system was then configured to enable task management and interrupt handling. To facilitate recovery and reestablishment, the LoRa connection frame counts must be maintained and synchronized with the network. This was achieved by saving the LMIC settings and MODEM frame counts in the nonvolatile memory area of the MCU before entering the deep-sleep mode. These were used on the wake-up and boot of the device to maintain the LoRa connection.

During the operation, following a wake-up or boot, the MCU performed the 125mS period of Leq(Z) audio measurement. The audio from the MEMS was sampled using the IIS module and DMA controller of the ESP32, which presents an audio block to be processed by the CPU at zero overhead. The audio block size was 6000 words long. This edge processing can be replaced by other algorithms, as required, and configured over the air. The prototype tested shows ample room for expansion with the current memory requirements as follows:

TABLE 1. Test data packets.

Date Byte	Description
0	Battery Level %
1	Annoyance level (0..255)
2	Leq(Z) in dB
3	Exposure Level (0..255)

RAM: 8.4% (used 27576 bytes from 327680 bytes)
Flash: 8.7% (used 290016 bytes from 3342336 bytes)

Following audio sampling and edge processing, the data produced were packetized into a four-byte data structure, as shown in Table 1.

As part of the LoRa communication, this data structure, along with other information (gateway status, device location, etc.), is forwarded via IP/UDP to the LoRaWAN network server. These data structures are then queried by an application server (MQTT is utilized in the prototype system) to form the user interface and allow remote access and monitoring.

Following successful message transfer, the node enters its deep-sleep mode. All power to the external peripherals is switched off using the GPIO of the MCU and FET combination before the MCU enters deep-sleep mode. The node is re-awakened at intervals set by the user using the MCU watchdog timer. In the test system, this was scheduled to occur at half-hourly intervals.

III. SYSTEM TESTING AND RESULTS

The focus of the testing was to establish the feasibility (in terms of accuracy and low-power operation) of providing a low-energy solution for the distributed audio measurements. The prototype was tested in two phases. Phase one was used to determine whether the MEMS had a measurement accuracy comparable to that of a more expensive sound-level meter-based solution. Phase two testing involved measuring the overall power consumption while testing the general and deep-sleep modes of operation.

A prototype test interface was built using TagIO [29], which demonstrated the possibility of remote access to a LoRaWAN network. This application can be executed remotely by using a wide range of devices.

A. POWER CONSUMPTION AND BATTERY USAGE

During testing, two nodes were deployed: one with a deep-sleep mode enabled, and the other that did not. Both nodes were booted and connected to the LoRaWAN network server, and the data packets were monitored and logged via the remote application interface.

The nodes were deployed in an urban setting within the kitchen of a house, approximately 5 miles (9 km) away from the concentrator/gateway. Both nodes were booted, and remote monitoring took place 18 miles (29 km) away, using a laptop and TagIO interface. The nodes were programmed to measure and send their data packets at half-hourly intervals. The distance of the monitoring station to the deployed nodes is largely irrelevant but is included as an example. The data packets sent by the nodes via the gateway are stored in the cloud, and are therefore accessible anywhere in the world at any time.

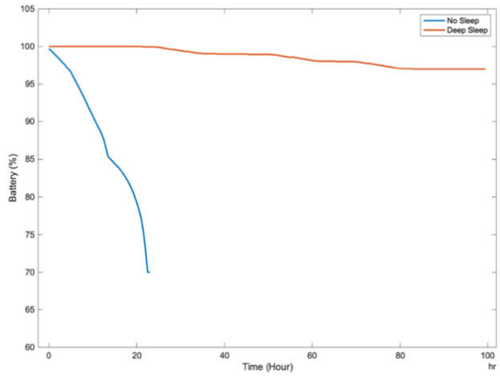

FIGURE 5. Continuous vs Deep Sleep operation time.

TABLE 2. Current consumption measurements.

Operation	Current Measure (mA)	Power Consumption
Continuous Audio Measurement (loop and display of audio levels to debug)	49.2	162mW
Single Audio Measurement and packet transmission and loop with no sleep	67	221mW
Single Audio Measurement and packet transmission followed by deep sleep mode	67 / 0.9 (deep sleep)	221mW / 3mW

Both nodes began with a 100% battery charge and were monitored to establish the duration of operation (Fig. 5a and 5b) using a 1000mAh battery. It was found that once the battery level dropped below 70%, the transmission began to fail. Therefore, a system was put in place to monitor this condition and inform the base station of the issue. This 70% level was used as a metric to compare the performances of the two nodes in terms of battery efficiency. Continuous current measurements were obtained from the deployed systems to establish deep and non-sleep consumption.

Current measurements were taken during the continuous operation of the nodes. All measurements were performed at the 1000mAh battery connection to the node. The results are presented in Table 2.

As can be observed in both Table 2 and Fig. 5 and 6, there is a significant improvement in the power efficiency and, therefore, the lifespan of the system. The power consumption reduces from 221mW to 3mW. The system with no deep-sleep implementation ran for 23 h, whereas the deep-sleep system ran for 761 h up to the 70% battery level.

Potentially, the deep-sleep system efficiency, could be improved even further by optimizing the circuit design. The development MCU board used during the testing, had an OLED display that, despite being off, may draw a small amount of current, likely via the connected IIC pins.

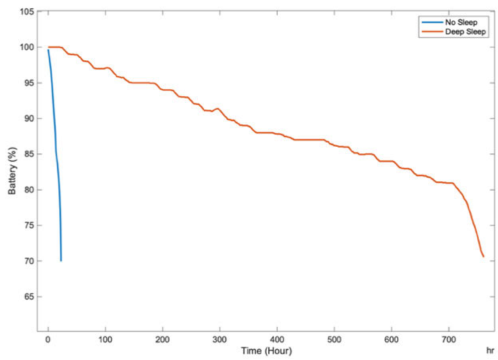

FIGURE 6. Continuous vs Deep Sleep operation time (larger time period).

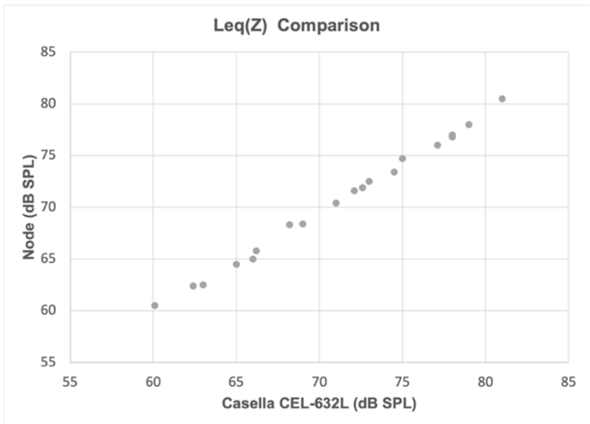

FIGURE 7. Node measurement vs calibrated Casella CEL-632L Noise Meter.

Further power savings can be achieved by optimizing the lower-power transmission modes of the LoRa network.

B. NODE MEASUREMENT ACCURACY

The node MEMs microphone accuracy was evaluated by comparing its measurements with those of a calibrated Casella noise level meter [38]. A pink noise source was used, and this was projected into the listening room using a Genelec 8040-powered speaker. A node was placed alongside the calibrated meter (1.25 meters from the speaker) to ensure measurements were taken in the far field. The measurements were performed in an ITU-R BS.1116 listening room [39].

Multiple measurements were obtained at points ranging between 60 dB (SPL) and 81 dB (SPL), resulting in an average SPL measurement available at each point for both the devices.

Fig. 7 illustrates that the node measurements offer a linear relationship with the calibrated Casella noise level meter, with a Pearson correlation coefficient of $r=0.998$. This indicates that despite its small size, the solution is highly accurate in performing these types of measurements. The *max, min, max(abs(difference))* and *RMSE* of the node measurements with respect to the Casella measurements were 1.2, 0, 1.2 and 0.513 dB respectively. This demonstrates the accuracy of the microphone sensitivity and therefore weighted measures (if properly implemented) would offer the same performance.

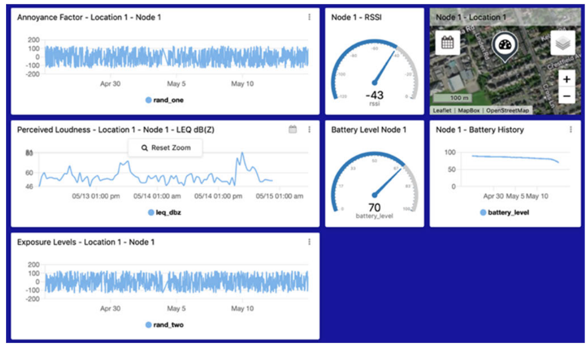

FIGURE 8. Test Interface showing Node 1 only.

C. EXAMPLE TEST INTERFACE

An example of the test interface is shown in Fig. 8. It indicates two random data bytes (Annoyance and Exposure Levels), the Leq(Z) calculated level, the node battery level, the location map, and the RSSI of node 1. This is the detected RSSI level from Node 1 to the gateway. The battery history was also presented.

The choice of whether to use a third-party or custom solution to create such an interface is largely based on the complexity required by the user. All the data packets, regardless of how they are displayed, are available via download for further processing. These data packets could be useful for noise prediction and, therefore, proactive reduction using machine-learning-based approaches.

IV. CONCLUSION

This paper discusses the importance of noise monitoring for health and well-being purposes and proposes a low-power and cost-effective solution that would enable distributed noise monitoring.

Using multilocation spatial monitoring, it is possible to study and predict how spatiality influences soundscape comfort or annoyance, thereby directly benefiting both urban sound management and school/hospital applications.

The results obtained indicate that a distributed noise monitoring solution can be achieved, which is accurate in its measurement, while offering low-power performance. The proposed solution was able to match the performance of a dedicated sound pressure meter, while the same time being significantly lower in cost and form factor. The system is also highly configurable, because edge processing can be adapted based on the application.

Successful transmission of Leq(Z), battery, annoyance, and location data was achieved with a minimal payload. This allows for easy incorporation of other metrics and data. These additional metrics allow for a comprehensive front-end application, thereby enabling predictive and preventative measures to be employed.

The direction of sound plays a role in annoyance; however, the traditional method of audio capture is using a single audio channel. Further studies utilizing spatial parameters could lead to the development of new annoyance models. For this purpose, a multichannel system is planned based upon the approaches outlined.

Additional work to identify appropriate measurement metrics and deployment strategies is also planned.

REFERENCES

[1] M. E. Beutel, C. Junger, E. M. Klein, P. Wild, K. Lackner, M. Blettner, H. Binder, M. Michal, J. Wiltink, E. Brähler, and T. Munzel, "Noise annoyance is associated with depression and anxiety in the general population—The contribution of aircraft noise," *PLoS ONE*, vol. 11, no. 5, May 2016, Art. no. e0155357, doi: 10.1371/journal.pone.0155357.

[2] D. Halperin, "Environmental noise and sleep disturbances: A threat to health?" *Sleep Sci.*, vol. 7, no. 4, pp. 209–212, Dec. 2014, doi: 10.1016/j.slsci.2014.11.003.

[3] O. Hahad, S. Kröller-Schön, A. Daiber, and T. Munzel, "The cardiovascular effects of noise," *Deutsches Ärzteblatt Int.*, vol. 116, no. 14, pp. 245–250, Apr. 2019, doi: 10.3238/arztebl.2019.0245.

[4] M. J. Jafari, R. Khosrowabadi, S. Khodakarim, and F. Mohammadian, "The effect of noise exposure on cognitive performance and brain activity patterns," *Open Access Macedonian J. Med. Sci.*, vol. 7, no. 17, pp. 2924–2931, Aug. 2019, doi: 10.3889/oamjms.2019.742.

[5] H. Ritchie and M. Roser. (Apr. 3, 2021). *Urbanization*. [Online]. Available: https://ourworldindata.org/urbanization accessed

[6] S. Kumar, K. Von Kriegstein, K. Friston, and T. D. Griffiths, "Features versus feelings: Dissociable representations of the acoustic features and valence of aversive sounds," *J. Neurosci.*, vol. 32, no. 41, pp. 14184–14192, Oct. 2012, doi: 10.1523/JNEUROSCI.1759-12.2012.

[7] N. P. M. Todd and F. W. Cody, "Vestibular responses to loud dance music: A physiological basis of the 'rock and roll threshol?"' *J. Acoust. Soc. Amer.*, vol. 107, no. 1, pp. 496–500, Jan. 2000, doi: 10.1121/1.428317.

[8] Y. Takeshima and J. Gyoba, "High-intensity sound increases the size of visually perceived objects," *Attention, Perception, Psychophys.*, vol. 75, no. 3, pp. 501–507, Apr. 2013, doi: 10.3758/s13414-012-0403-z.

[9] *LoRa, What is LoRa*. Accessed: May 26, 2021. [Online]. Available: https://lora-developers.semtech.com/get-started/what-is-lora/

[10] E. U. Directive, "Directive 2002/49/EC of the European parliament and the council of 25 June 2002 relating to the assessment and management of environmental noise," *Off. J. Eur. Communities, L*, vol. 189, pp. 11–14, Jul. 2002.

[11] *Acoustics—Description and Measurement of Environmental Noise—Part 2: Acquisition of Data Pertinent to Land Use*, Standard ISO 1996-2:1987, 1987.

[12] *Acoustics—Description, Measurement and Assessment of Environmental Noise—Part 2: Determination of Sound Pressure Levels*, Standard ISO 1996-2:2017, 2017.

[13] V. G. Escobar, J. M. B. Morillas, G. R. Gozalo, R. Vílchez-Gómez, J. C. D. Río, and J. A. M. Sierra, "Analysis of the grid sampling method for noise mapping," *Arch. Acoust.*, vol. 37, no. 4, pp. 499–514, Dec. 2012, doi: 10.2478/v10168-012-0062-z.

[14] *Wireless Sound Level Meter/Data Logger, Convergence Instruments, NSRTW_mk3*. Accessed: May 18, 2021. [Online]. Available: https://convergenceinstruments.com/product/wireless-sound-level-meter-data-logger-nsrtw_mk3/

[15] B. Shield and J. E. Dockrell, "External and internal noise surveys of London primary schools," *J. Acoust. Soc. Amer.*, vol. 115, no. 2, pp. 730–738, Feb. 2004, doi: 10.1121/1.1635837.

[16] J. L. Darbyshire and J. D. Young, "An investigation of sound levels on intensive care units with reference to the WHO guidelines," *Crit. Care*, vol. 17, no. 5, p. R187, 2013, doi: 10.1186/cc12870.

[17] C. Roberts and E. Schweizer, "Noise monitoring: A strategy for better patient care in hospitals," Schneider Electr., U.K., White Paper 998-2095-04-09-15BR0_EN, 2015.

[18] M. Vacher, "Smart audio sensor for telemedicine," in *Proc. Smart Object Conf.*, vol. 225, May 2003, pp. 222–225.

[19] M. A. Shah, I. A. Shah, D.-G. Lee, and S. Hur, "Design approaches of MEMS microphones for enhanced performance," *J. Sensors*, vol. 2019, pp. 1–26, Mar. 2019, doi: 10.1155/2019/9294528.

[20] C. P. Chen, C. L. Chuang, and J. A. Jiang, "Ecological monitoring using wireless sensor networks—Overview, challenges, and opportunities," in *Advancement in Sensing Technology*, vol. 1. Berlin, Germany: Springer, 2013, doi: 10.1007/978-3-642-32180-1_1.

[21] A. Smeaton and M. McHugh, "Towards event detection in an audio based sensor network," in *Proc. 3rd ACM Int. Workshop Video Surveill. Sensor Netw.*, Nov. 2005, pp. 87–94.

[22] R. Mangharam, A. Rowe, R. Rajkumar, and R. Suzuki, "Voice over sensor networks," in *Proc. 27th IEEE Int. Real-Time Syst. Symp. (RTSS)*, Dec. 2006, pp. 291–302.

[23] V. Berisha, H. Kwon, and A. Spanias, "Real-time acoustic monitoring using wireless sensor motes," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2006, pp. 847–850.

[24] A. Harma, M. F. McKinney, and J. Skowronek, "Automatic surveillance of the acoustic activity in our living environment," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2005, pp. 1–4.

[25] K. V. V. Girish, A. G. Ramakrsihnan, and N. Kumar, "A system for distributed audio classification using sparse representation over cloud for IoT," in *Proc. 10th Int. Conf. Commun. Syst. Netw. (COMSNETS)*, Jan. 2018, pp. 342–347, doi: 10.1109/COMSNETS.2018.8328217.

[26] S. Kumari, D. Roy, M. Cartwright, J. P. Bello, and A. Arora, "EdgeL^3: Compressing L^3-Net for mote scale urban noise monitoring," in *Proc. IEEE Int. Parallel Distrib. Process. Symp. Workshops (IPDPSW)*, May 2019, pp. 877–884.

[27] *IEEE Standard for Information Technology—Telecommunications and Information Exchange between Systems Local and Metropolitan Area Networks—Specific Requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 1: Enhancements for High-Efficiency WLAN*, Standard IEEE 802.11ax-2021, May 19, 2021.

[28] *Bluetooth Core Specification*, BlueTooth SIG, Kirkland, WA, USA, V5.2, Dec. 31, 2019.

[29] *TagoIO*. Accessed: May 21, 2021. [Online]. Available: https://tago.io

[30] *The Things Network (TNN)*. Accessed: May 21, 2021. [Online]. Available: https://www.thethingsnetwork.org/support

[31] *Invensense MEMs Mic*. Accessed: May 21, 2021. [Online]. Available: https://invensense.tdk.com/wp-content/uploads/2015/02/INMP441.pdf

[32] *Inter-IC-Sound Specification (I2S)*. Accessed: May 21, 2021. [Online]. Available: https://www.cypress.com/file/133906/download

[33] *ESP32 Reference Manual, Espressif Systems*. Accessed: May 21, 2021. [Online]. Available: https://www.espressif.com/sites/default/files/documentation/esp32_technical_reference_manual_en.pdf

[34] *Semtech LoRa MODEM, Sx127x*. Accessed: May 21, 2021. [Online]. Available: https://semtech.my.salesforce.com/sfc/p/#E0000000JelG/a/2R0000001Rbr/6EfVZUorrpoKFfvaF_Fkpgp5kzjiNyiAbqcpqh9qSjE

[35] *Free Real Time Operating System*. Accessed: May 21, 2021. [Online]. Available: https://www.freertos.org

[36] *IBM LoRaWAN in C Port*. Accessed: May 21, 2021. [Online]. Available: https://github.com/mcci-catena/arduino-lmic

[37] A. N. Rimell, N. J. Mansfield, and G. S. Paddan, "Design of digital filters for frequency weightings (A and C) required for risk assessments of workers exposed to noise," *Ind. Health*, vol. 53, no. 1, pp. 21–27, 2015, doi: 10.2486/indhealth.2013-0003.

[38] *Casella Noise Meter*. Accessed: Jun. 20, 2021. [Online]. Available: https://www.casellasolutions.com/uk/en.html

[39] *International Telecommunication Union 2015 Recommendation ITU-R BS.1116-3: Methods for the Subjective Assessment of Small Impairments in Audio Systems*. Accessed: May 21, 2021. [Online]. Available: https://www.itu.int/rec/R-REC-BS.1116/en

STEVEN FENTON received the bachelor's degree in electronic and information engineering, in 1994, and the Ph.D. degree, in 2017, for his work on the proposal of new perceptual-based audio measurements.

Since 2007, he has been working as a Senior Lecturer and the Course Leader of the University of Huddersfield, U.K. He is currently a Senior Research and Development Engineer, he developed cutting-edge technology within the professional and consumer marketplace. This work included video and audio compression systems, in-car audio, broadcast mixing console development, digital and analogue audio design, algorithm, low power embedded development, and project management.

Dr. Fenton is an Active Member of the AES and a fellow of the Higher Education Academy. He is also a member of the Applied Psychoacoustics Laboratory, University of Huddersfield. His research interests include the measurement of audio quality in music production, audio mastering, broadcast, automatic mixing, and artificial intelligent systems.