IoT Applications with Common Distributed Architecture for Data Acquisition

Kushal Thapa
Ingram School of Engineering
Texas State University
San Marcos, USA
email: k_t260@txstate.edu

Vinay Lokesh
Ingram School of Engineering
Texas State University
San Marcos, USA
email: v_v183@txstate.edu

Kevin Seets
Ingram School of Engineering
Texas State University
San Marcos, USA
email: kms489@txstate.edu

***Abstract*—Internet of Things (IoT) applications have many forms, and leverage many different technologies. However, certain classes of applications have extremely strong similarities in system architecture. This paper discusses several important applications, which leverage a small set of loosely coupled, distributed data acquisition subsystems to effect a centrally coordinated, data intensive function or application. For brevity, we call this structure "Coordinated IoT for Data Acquisition" (CIDAQ). Finally, this paper introduces a novel power line communication research, which employs this CIDAQ architecture for data capture and processing.**

Keywords—Internet of Things; IoT; Smart Grid; Active Shooter; Machine Learning; ML; Artificial Intelligence; AI.

I. INTRODUCTION

In some scenarios where Internet of Things (IoT) technologies are used, the application of interest could be useful for training first responders, as it provides a more analytical approach to how first repsonders move and react. In other cases, IoT's use relates more to optimization of an industrial process, where a net of data could increase functionality or longevity. It could also have use in observations of biological processes where data has proven difficult to gather by more traditional means. In all cases, the application benefits from the Coordinated IoT for Data Acquisition (CIDAQ) architecture, where a distributed, loosely-coupled set of IoT devices provides telemetry data to a central repository, and Machine Learning and Artificial Intelligence (ML/AI) algorithms are employed to produce some application-related insights.

In this paper we describe various applications which leverage the CIDAQ architecture as well as some useful technologies. Section II presents a compelling application related to training of first responders in active shooter scenarios. Section III briefly describes several other applications, and presents some of the ML/AI techniques that can be useful in these applications. Section III-D exhibits some businesses and products that are already available for purchase, ranging in size and scope.

We conclude the paper in Section IV with a particularly interesting application of the CIDAQ architecture, which focuses on the electrical grid. In this application, the distributed system "listens" to current disturbances on the electrical distribution grid, "talks" upstream from the outlet to the substation, and "geolocates" electrical devices for system management purposes based on actively and passively gathered telemetry data.

II. ACTIVE SHOOTER TRAINING

A particularly compelling application of IoT systems is to augment the training of first responders. This application leverages the CIDAQ architecture by placing data acquisition devices and sensors on various body parts of first responders who are participating in scenario-based, real-time active shooter training.

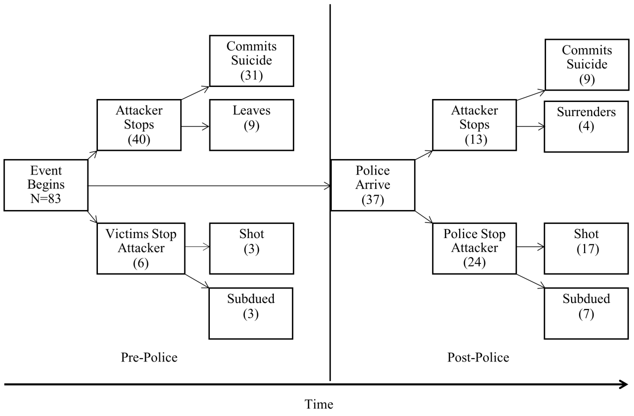

Fig. 1. Comparison of active shooter situations with and without first responder intervention [1].

Having properly trained first responders immediately available for active shooter situations is paramount to ensuring the safety and survival of bystanders. Fig. 1 indicates an increase in victim shootings and death associated with police intervention. In a comparison study of 83 events with and without police intervention, 37 (45%) included police intervention, which accounted for over half of total victim shootings (63%) and deaths (56%) [1]. Though this is a particularly alarming trend, it does give insight on the fact that police officers need to be better trained to deal with active shooter situations.

Traditional training for an active shooter situation can be time consuming and expensive, and so is often not effective in producing measurable outcomes [2]. The prevalence of active shooter situations in the United States reveals a necessity for officer training that is effective in measurable outcomes as well as conventional metrics of time and cost [3]. Students at Texas State University have designed a system with the potential to improve training for first responders which could play a part in revolutionizing first responder training, as shown in Fig. 2. This system leverages the CIDAQ architecture. By placing sensors strategically on participant, data about proper movement and weapon handling can be gathered in real-time from multiple participants, processed, and analyzed in a central location, and leveraged to improve the effectiveness of first-responder training. This information can in turn be used to create augmented reality training programs that could be effective tools in saving law enforcement valuable time and money, and in improving the ability to repeat training remotely [2]. Additionally, the data can be used to precisely compare

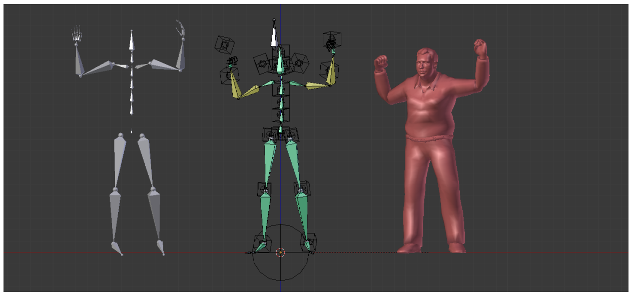

Fig. 2. Augmented motion-tracking of first responders using CIDAQ architecture sensors [4].

and contrast the effectiveness of different training programs, thereby improving the effectiveness of the training program as well as the measurable outcomes for each participant.

The devices used in this application consist of Inertial Measurement Units (IMU) placed on the head, chest, and weapon of the participants. Each IMU contains analog orientation, acceleration, and location sensors whose outputs can be easily acquired, stored, measured, and manipulated. The data collected by these devices creates a 3-dimensional map of the user's movements which can be reconstructed and replayed at-will. Augmented reality devices can create virtual training experiences with these maps that include all of the necessary movements and actions required by a first responder [3]. This training method not only improves time and cost effectiveness, but also affords a new level of access. The ease of distribution of these expert training programs increases accessibility for smaller or more distant municipalities and also helps streamline training. This ensures first responders have access to the exact same training, which will help with joint operations as well as transfers between departments.

The application of CIDAQ architecture for training first responders does not stop at active shooter situations; there are a number of other projects being developed and at least one that is already in use. Dartmouth College's Interactive Media Laboratory and Institute for Security Technology Studies created a virtual program designed to aid in training for terrorism response. The program, called Ops-Plus, utilizes 3D simulators to aid in training against attacks that involve nuclear, radiological, biological, and even chemical warfare. The Los Angeles Police Department currently uses an immersive simulation trainer, called HYDRA, used to train first responders in a series of scenarios, from earthquakes to terrorism response, that would be difficult to recreate. New York City has partnered with the Environment Tectonics Corporation to develop a software, similar to HYDRA, that creates an immersive environment designed to help first responders prepare for citywide disaster management [5]. As these technologies progress, the breadth and depth of IoT application will increase, reducing potential harm to first responders and civilians.

III. SIMILAR APPLICATIONS AND ML/AI

In addition to applications such as training for first responders, the CIDAQ architecture is being actively deployed in various other applications including monitoring and control of industrial processes, monitoring of the habitat for endangered species, and enabling efficient hospital care for bedridden patients. These applications are described briefly along with the general nature of ML/AI algorithms which could be used in processing the resulting telemetry data to create new knowledge, or to improve application-specific outcomes.

A. *Industrial Application*

Heat trace cables or heat tapes are vital in oil and gas, chemical treatment, power generation and many other industries. These cables and their control systems assist in the continuous delivery of gases and liquids, often preventing the contents of pipes or tanks from freezing in extreme environments.

Based on advances in heat sensitive polymer design, many of these cables come with self-regulating ability. In other words, the heat generated by these cables can compensate for environmental temperatures by autonomously adapting their absorption of electrical current [6]. This capability provides operational simplicity in external power control systems, as well as convenience in direct attachment to the electrical power source.

However, the heat cable can be damaged during deployment, or can degrade due to aging or other conditions. Approaches to locating damaged portions of heat cable is a difficult challenge. One approach to this remote monitoring problem is to integrate temperature sensors into the cable, or add IoT-based devices along the cable to make measurements. As a distributed, network-based monitoring architecture, a CIDAQ system is a logical candidate for this application. Via a CIDAQ architecture, deployed heat cables and the systems they monitor can be remotely evaluated using low-rate data transmitted along the heat cable and power lines. This data, transmitted directly via low-frequency power line communications techniques, can also be aggregated, assimilated, and analyzed using the ML/AI algorithms.

B. *Biodiversity Application*

The conservation of endangered species is important for maintaining biodiversity and a well-balanced ecosystem. Several techniques have only recently been applied in the marine environment to detect the presence of marine species [7]. Confirming presence relies on locating the animals, which can prove challenging for species with low population numbers. A variety of methods have been used to determine the presence of rare marine species, including fishing and underwater visual surveys [8]. However, these approaches typically require substantial field-based effort by researchers and data gatherers. Although scientists have been able to achieve a significant amount of success using Environmental DNA (eDNA), not every organism will be readily detected by eDNA and the scale of the water bodies impacts the probability of detection.

Importantly, the abiotic factors of temperature, UV radiation, and amount of DNA present all impact the length of time that the eDNA stays in the environment [9]. A compelling approach which uses the CIDAQ architecture to monitor endangered species is detects and processes animal voice or audio signals. Examples of such endangered species that are being monitored with a CIDAQ architecture are the Houston Toad and Craw Frog [10]. In these applications, an embedded solution detects toad calls automatically with real-time notification transmission capabilities to engage remote researchers. The labelled audio data is filtered and fed to a machine learning model to extract features. The extracted features are then fed to classification algorithms using the processing pipeline shown in Fig. 3.

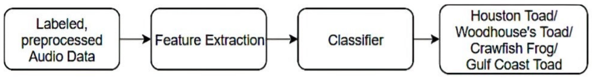

Fig. 3. Experimental Method.

This application of the CIDAQ architecture leverages deep learning architectures, such as Recurrent Neural Network (RNN), Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), Gated Recurrent Units (GRUs) as well as conventional signal processing such as Mel-frequency Cepstral Coefficients (MFCC), Linear predictive coding (LPC), Perceptual Linear Prediction (PLP), Mel Filter banks, and Spectrograms [11]. These technologies are used to improve the identification of endangered species with reduced false-positive rate [10].

C. Medical Application

Another intuitive application that uses CIDAQ architecture enables efficient hospital care for bedridden patients. The feasibility of using pervasive sensing technology and artificial intelligence for autonomous and granular monitoring in the Intensive Care Unit (ICU) is vital since manual observations can suffer from subjectivity. The use of sensing technologies and network-based telemetry can bring timely intervention to assist in making life-saving decisions while dealing with high levels of uncertainty under strict time constraints [12]. Artificial intelligence in the critical care unit could reduce doctors' workload to allow them to spend time on more critical tasks. The approach used in this application include accelerometer sensors, a light sensor, a sound sensor, and a high-resolution camera to capture data on patients and their environment in the ICU. Various computer vision and deep learning techniques are used to recognize a patient's face, posture, facial expressions, head pose, and extremity movements from video data [13]. For activity recognition, data from wearable accelerometer sensors worn on the wrist, ankle, and arm are analyzed. Additionally, the information uses the room's sound pressure and light intensity levels to examine their effect on patients' sleep quality. This framework employs a cascaded architecture with three stages of deep CNNs to predict face and landmark locations in a coarse-to-fine manner [12].

In general, most embedded applications dealing with IoT, machine learning and artificial intelligence implement a CIDAQ architecture. Many of these applications leverage deep learning algorithms, which are concerned with very large datasets of labelled analog data, such as image, text, audio and video [14]. Machine learning algorithms used in CIDAQ-based systems can be based on supervised, unsupervised or semi-supervised learning. Supervised learning model is based on training data and helps to make predictions, some of the important algorithms under supervised learning are logistic regression and back propagation neural networks. Unsupervised learning model is prepared by deducing structures present in the input data, some of the important algorithms under unsupervised learning are Apriori and K-means algorithm. Semi-supervised learning is a mixture of both labelled and unlabeled data.

D. Market Data

As IoT technologies gain in popularity and scope of capability they will become more available and more widely used. There are a number of businesses–ranging from startups to Fortune 50 companies–using IoT, and implementing CIDAQ architecture in a range of ways, for their products and services. A few smaller startup companies are: Vicotee, MachineMax, and Radio Bridge [15]. Vicotee is a Norwegian company that offers a variety of sensors that can be used in conjunction with each other for myriad 'smart' applications including but not limited to shipping, infrastructure, healthcare, air quality, and land management [16]. MachineMax, based out of the UK, on the other hand offers a software, rather than only offering a full-package, that uses IoT to interconnect 3rd party sensors into a a single platform [17]. Similarly, US based Balena offers IoT 'fleet management' that is designed to push updates across varying platforms using a cloud-based container [18].Instead of simply offering a product line, US based Radio Bridge offers IoT data-as-a-service where they set up devices and monitor the data, allowing the end user to focus on whatever task is at hand [19]. Small startups are not the only businesses interested in IoT technologies, there are also several major players competing in the market as well, specifically IBM, Samsung, AWS, and Microsoft [20]. Each of these Global 500 businesses offers a proprietary monitoring solution for their products.

IV. ELECTRIC DISTRIBUTION GRID

As mentioned in Section I, the electric distribution grid is a particularly compelling application of the CIDAQ architecture. As more compute and sensor devices pervade modern society, reliance on the electrical infrastructure continues to increase. Although "the grid" is one of the unspoken wonders of the modern technological society, the increased burdens of two-way power flow, distributed generation facilities, and complex/dynamic structure are pressing technologists to create better approaches to monitoring, controlling, and leveraging existing grid infrastructure.

An intuitive approach to leverage existing grid infrastructure is to use the grid itself as a communication medium. This would pave a way for the development of a self-regulating power grid, and in addition, obviate the need for deploying other communication resources for grid related applications, thereby saving time and money worth billions [21]. This distributed system concept fits precisely within the CIDAQ architecture.

Using the power grid as a communication medium is not a novel concept. This technology, commonly referred to as Power Line Communication (PLC), has been used since early 1920's for applications like fault detection and automatic meter reading [22]. However, the development in the PLC applications has been severely hindered due to the dynamic and unpredictably noisy nature of this medium. This problem is aggravated by the different electrical devices in the power grid, like transformers, which obstruct and muddle the communication signals even more [23].

One simple approach to solving this problem is the transmission of low-frequency communication signals. Low- frequency signals don't attenuate or distort as much compared to high-frequency signals, even after passing through the transformers [24]. Thus, a low-frequency band, typically in the range of 150 Hz-1350 Hz [24], can be used directly for PLC applications. However, one major disadvantage of using low- frequency bands for communication is the low data rate. Consequently, this solution has been under-researched and

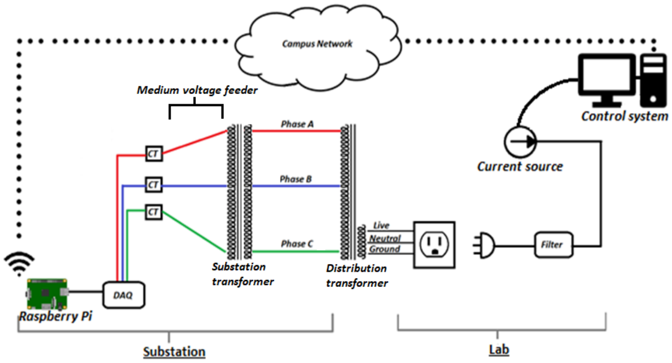

Fig. 4. Experimental setup for sending and capturing communication signals through powerlines.

mostly overlooked.

Nonetheless, such low-frequency PLC has applications in fields that don't require high-speed data transmission but prioritize reliability, simplicity, scalability, and ease of deployment. There is an urgent need for this type of technology in the power sector. As mentioned before, the existing power grids are failing because of the exponential increase in power demand over the last few decades [25]. This problem is exacerbated by the disconnect between the power producers and consumers which leads to a huge waste of already depleted power supply. Low-frequency PLC can bridge this disconnect thereby becoming the backbone communication infrastructure of a continuously sensing and self-monitoring power grid called "smart grid" [26].

Therefore, there is a need for research and study in the field of low-frequency PLC. To that end, our research is going to be focused on employing a CIDAQ network architecture to test low-frequency power line communication for simple digital communication. Fig. 4 shows a simplified design of this network architecture.

As shown in Fig. 4, a programmable current source injects a known signal into the powerline via a stabilizing filter. This signal passes through single-phase and three-phase distribution grid via transformers. When the signal passes through a transformer, the image of the signals are ingrained in the other two-phases of the power line, creating noisy image or echo signals [23]. In the electric substation, the injected signals, which have passed through multiple transformations, are collected by a Data Acquisition Device (DAQ) [27]. The DAQ is controlled by an embedded system such as a Raspberry Pi [28]. The control system and remote DAQ devices are connected to a common wireless network for synchronization. The control system commands the programmable current source to inject communication or disturbance signals, and simultaneously commands the DAQ to acquire signal data. This data is collected centrally for signal processing and machine learning techniques to analyze and reconstruct the original signal, remove extraneous noise or images of signals, and create a global understanding of the signal context on the distribution grid.

The raw data captured at the substation contains the input communication signal mixed with a more dominant power signal and its highly correlated harmonics, plus a highly dynamic and unpredictable environmental noise. Using the CIDAQ architecture, the distributed system extracts the information sent by the input signal from this mixture. Traditional Digital Signal Processing (DSP) techniques [29] do this job in other communication media like telephone wires and optical fibers, but are not adequate for highly noisy PLC. As a result, supplementing DSP techniques with ML becomes a critical aspect of this CIDAQ implementation. Machine learning is a data-driven technique that formulates a relationship between the input and the output based on the training data [30]. In the case of the ultra-low-frequency powerline environment, the training data is composed of various features of the raw signal extracted using DSP techniques. This data is fed into ML algorithms to form a model which extracts the transmitted information from the raw noisy data. The overall signal flow from transmitted input to the extracted ML output is shown in Fig. 5.

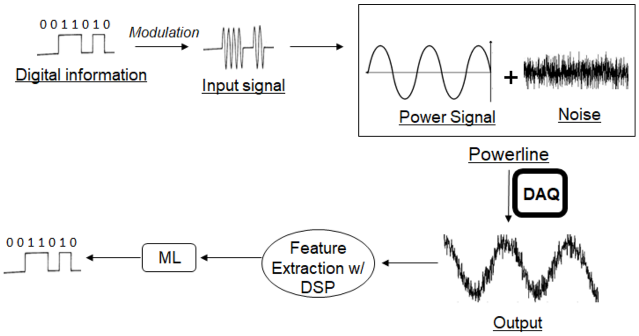

Fig. 5. Flowchart showing the basic signal workflow from input to machine learning output.

This way, the CIDAQ architecture effectively captures PLC data from the electrical grid and deciphers the information contained in the data. This CIDAQ architecture can also easily be scaled to employ multiple DAQs, which can be placed at differing locations in the power grid. A similar architecture can also be used to capture non-PLC internal power grid data, which can contain information about the state of the grid and its components. This non-PLC data can similarly be processed and deciphered using signal processing and ML techniques via the same CIDAQ architecture.

V. CONCLUSION

Using the context of several related applications, this paper has introduced the CIDAQ architecture, which leverages distributed IoT-like devices to acquire relatively high-rate signals and analyze them collectively to produce some application-specific outcome. In cases of first responder training, IoT devices are distributed on a participant's body for analysis of form and function in a high intensity situation. In cases of industrial control and management, IoT devices are distributed on heat-tape which ensures the consistent flow of gas or liquid in an industrial environment. In cases of species management or healthcare management, IoT devices gather data from the environment or from healthcare facilities for processing and analysis, and to create a larger context from which to prioritize societal or personal decisions.

In the electrical grid, which underpins almost all related applications, sensors, and data acquisition systems are distributed throughout a larger, dynamic context in order to acquire signals, gather data, cross-correlate events, and effect changes in system efficiency that will enable future applications which could benefit from the CIDAQ architecture.

ACKNOWLEDGEMENT

The authors would like to thank Bo Hyse, Matt Healea, and Karina Paz for the use of their senior design project as an example in this paper, as well as the Advanced Law Enforcement Rapid Response Training (ALERRT) team at Texas State University for the use of their research involving active shooter training.

REFERENCES

[1] M. H. Martaindale, W. L. Sandel, and J. P. Blair, "Active-shooter events in the workplace: Findings and policy implications," *J. Bus. Contin. Emer. Plan.*, vol. 11, no. 1, p. 6–20, 2017.

[2] P. Glass *et al.*, "Use of computer simulation modeling to reduce the consequences of an active shooter event in a large event venue," in *2018 IEEE International Symposium on Technologies for Homeland Security (HST)*, 2018, pp. 1062–1065.

[3] B. Heyse, M. Healea, and K. Paz, "EE Senior Design project Statement of Work: 3-D Motion Capture and Virtualization for Response Training," Ingram School of Engineering, Texas State Univ., San Marcos, Texas, Senior Design Thesis, 2019.

[4] Miecz. Blender motion capture retargeting tip. 25-March-2021. [Online]. Available: https://www.blendernation.com/2018/08/29/blender-motion-capture- retargeting-tip/

[5] G. Meyers. Simulation software is cost-effective for emergency response training. FyorgeFX Training Simulations. 25-March-2021. [Online]. Available: https://forgefx.com/simulation-software-is-cost-effective-for-emergency-response-training/

[6] USX^{TM} Self-Regulating Heating Cable. Thermon. 25-March-2021. [Online]. Available: http://www.thermon.com/us/products/electric-heating/heat-trace/usx#overview

[7] K. Weltz *et al.*, "Application of environmental DNA to detect an endangered marine skate species in the wild," *PLoS ONE*, vol. 12, pp. 1–16, June 2017.

[8] G. F. Ficetola *et al.*, "Species detection using environmental DNA from water samples," *Biology letters*, vol. 4, pp. 423–5, Sep. 2008.

[9] H. C. Rees *et al.*, "The detection of aquatic animal species using environmental DNA – a review of eDNA as a survey tool in ecology," *Journal of Applied Ecology*, vol. 51, pp. 1450–1459, 2014. [Online]. Available: http://eprints.nottingham.ac.uk/id/eprint/30254

[10] S. Islam and D. Valles, "Houston Toad and other chorusing amphibian species call detection using deep learning architectures," in *Proc. 10th IEEE Computing and Comm. Workshop and Conf.*, Jan. 2020, pp. 0511–0516.

[11] L. Rabiner and R. Schafer, *Theory and Applications of Digital Speech Processing*. Pearson, 2011.

[12] A. Davoudi *et al.*, "Intelligent ICU for Autonomous Patient Monitoring Using Pervasive Sensing and Deep Learning," *Scientific Reports*, vol. 9, pp. 1–13, May 2019.

[13] N. A. Halpern and S. M. Pastores, "Critical care medicine in the united states 2000-2005: an analysis of bed numbers, occupancy rates, payer mix, and costs," *Critical care medicine*, vol. 38, no. 1, p. 65—71, Jan. 2010. [Online]. Available: https://doi.org/10.1097/CCM.0b013e3181b090d0

[14] Deep Learning. Machine Learning. 25-March-2021. [Online]. Available: https://machinelearningmastery.com/what-is-deep-learning/

[15] 5 top wireless sensor solutions impacting the telecom sector. 30-March-2021. [Online]. Available: https://www.startus-insights.com/innovators-guide/5-top-wireless-sensor-solutions-impac ting-the-telecom-sector/

[16] How vicotee's technology can help you move fast forward. 30-March-2021. [Online]. Available: https://https://www.vicotee.com/use-cases/

[17] We're building the new age of construction. 30-March-2021. [Online]. Available: https://machinemax.com/pages/solutions

[18] What is balena? 08-April-2021. [Online]. Available: https://www.balena.io/what-is-balena

[19] Iot data-as-a-service. 30-March-2021. [Online]. Available: https://radiobridge.com/services/iot-data-as-a-service

[20] J. Guth *et al.*, *A Detailed Analysis of IoT Platform Architectures: Concepts, Similarities, and Differences*. Springer, 2018, pp. 81–101.

[21] K. Rabie *et al.*, "IEEE Access Special Section Editorial: Advances in Power Line Communication and its Applications," *IEEE Access*, vol. 7, pp. 133 371–133 374, Sep. 2019.

[22] K. Dostert, "Telecommunications over the power distribution grid - possibilities and limitations," in *Proc. Int. Symp Power Line Comm. and Appl.*, 1997, pp. 1–9.

[23] S. U. Ercan *et al.*, "Power line communication design and implementation over distribution transformers," in *2017 10th International Conference on Electrical and Electronics Engineering (ELECO)*, 2017, pp. 190–194.

[24] D. Dzung, I. Berganza, and A. Sendin, "Evolution of powerline communications for smart distribution: From ripple control to OFDM," in *2011 IEEE International Symposium on Power Line Communications and Its Applications*, 2011, pp. 474–478.

[25] S. McClellan, D. Valles, and G. Koutitas, "Dynamic voltage optimization based on in-band sensors and machine learning," *Appl. Sci.*, vol. 9, no. 14, p. 2902, July 2019.

[26] N. Uribe-Pérez *et al.*, "Smart grid applications for a practical implementation of IP over narrowband power line communications," *Energies*, vol. 10, no. 11, p. 1782, Nov. 2017.

[27] DI-1100 4-channel USB Data Acquisition Starter Kit. DATAQ Instruments. 25-March-2021. [Online]. Available: https://www.dataq.com/products/di-1100/

[28] Raspberry Pi 4 Model B. Raspberry Pi. 25-March-2021. [Online]. Available: https://www.raspberrypi.org/products/raspberry-pi-4-model-b/

[29] D. Kakati and S. C. Arya, "A full-duplex optical fiber/wireless coherent communication system with digital signal processing at the receiver," *Optik*, vol. 171, pp. 190 – 199, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S003040261830799X

[30] S. Raschka and V. Mirjalili, *Python Machine Learning*, 2nd ed. Packt Publishing, 2017.