Design, Application, and Validation of an IoT Wireless Sensor Network based on Lora for Strawberry Farming

Mateus R. da Cruz and Samuel Mafra and Felipe A. P. de Figueiredo

***Abstract*— Agriculture innovation has become essential to increasing the efficacy of food production. Wireless Sensor Networks and the Internet of Things can play an essential role in this innovation process. This paper proposes and validates a LoRa network that applies Machine Learning capabilities on edge following an Internet of Things approach. The Wireless Sensor Network proposed combines Long Range communications and monitoring systems, allowing a data-driven farming approach. It is essential in a world with a crescent population and limited plantation area.**

***Keywords*— Internet of Things, LoRa, Machine Learning, Wireless Sensors Network.**

I. INTRODUCTION

Internet of Things (IoT) applications have recently been intensely developed and introduced to agriculture. IoT solutions are being proposed for greenhouses, livestock, and field monitoring. These applications use Wireless Sensors Networks (WSN) to connect devices and exchange real-time data. Solutions following a data-driven approach are extensively developed for the crop and livestock monitoring [1]. Each crop solution has multiple subdomains based on the variable of interest like soil, air, temperature, water, and disease monitoring [2]. In contrast, many subdomains can be implemented in an all-in-one solution, providing flexibility to work in more scenarios and situations. This platform approach is used in the proposed solution, enabling the simultaneous use of several services.

By the characteristics of IoT, the main challenges in developing IoT applications are to minimize the power consumption and maximize the coverage of the networks. The implementation of IoT in the agriculture domain has multiple applications, protocols, and prototypes that must be aggregated [1]. Some offer necessary characteristics for IoT devices, like low-power consumption and low hardware complexity. For the communication with long coverage in rural areas, the Long Range (LoRa) has been proposed as an ideal solution. LoRa is a radio-frequency technology that offers long-range communication, low-power consumption, and flexibility [3]. These characteristics are essential for IoT agriculture applications and enable extended area coverage.

Also, cheap and affordable devices have been widely developed in recent years. For example, Raspberry Pi and Arduino boards are the most common prototyping boards for developing IoT applications. These boards support Machine Learning (ML), Computer Vision (CV), automation, and more technologies. This paper proposes a WSN using LoRa to apply data gathering and ML to strawberry farming. The application construction aims to reduce implementation costs and increase the scalability of application in large plantations. Also, the proposed solution uses ML capabilities to present a better irrigation system.

The remainder of this paper is organized as follows. Section II presents a brief literature review and some technologies used in IoT agriculture solutions. Section III describes the proposed IoT application architecture and presents a detailed workflow of the services provided. Section IV presents and analyzes the results. Finally, Section V concludes the paper and suggests further future works.

II. RELATED WORK

Research on IoT agriculture applications has led to several approaches for capturing and processing crop data. As a result, IoT solutions for agriculture have been widely developed to increase the efficiency, the quality of products, and production of farms. Also, most of these applications aim to be sustainable and reduce the environmental impacts caused by some processes in agriculture [4].

Most of these applications focus on monitoring and controlling systems. Monitoring solutions use widely distributed sensors on the farm to monitor critical variables like humidity and soil moisture in real-time. In addition, some offer cloud-based services to analyze the collected data and give crop information, with the aim of automating farming processes like irrigation and temperature in greenhouses [2]. In the same way, applications based on positioning, tracking, and monitoring are being developed for tracking livestock and detecting illnesses. This approach is necessary as the number of livestock grows, as well as the production costs [5].

IoT solutions have already been proposed to monitor critical variables that affect plant growth and production. Also, LoRa

Mateus Raimundo da Cruz, Instituto Nacional de Telecomunicações, Santa Rita do Sapucaí-MG, e-mail: mateusraimundo@mtel.inatel.br; Samuel Baraldi Mafra, Departamento de Engenharia de Telecomunicações, Instituto Nacional de Telecomunicações, Santa Rita do Sapucaí-MG, e-mail: samuelb-mafra@inatel.br; Felipe Augusto Pereira de Figueiredo, Departamento de Engenharia de Telecomunicações, Instituto Nacional de Telecomunicações, Santa Rita do Sapucaí-MG, e-mail: felipe.figueiredo@inatel.br. This work is partially supported by RNP, with resources from MCTIC, Grant No. No 01245.010604/2020-14, under the Brazil 6G project of the Radiocommunication Reference Center (Centro de Referência em Radiocomunicações - CRR) of the National Institute of Telecommunications (Instituto Nacional de Telecomunicações - Inatel), Brazil, the National Council for Scientific and Technological Development-CNPq (403827/2021-3), FAPESP (2021/06946-0)

communication based on LoRaWAN protocol has already been develop and tested in greenhouses applications following a Farming as a Service (FaaS) approach [6]. In [4], an IoT platform that controls an irrigation system is proposed. The platform uses the LoRaWAN network protocol and a WSN approach to collect and deliver critical information to the user. Also, the platforms offer other services such as the alarm of possible pest events and monitoring solutions. In a brief, the IoT platform presents ways to control the irrigation system more effectively, saving energy and water.

Many technologies, protocols, and devices are integrated into a single proposal for these applications to become possible. For example, some solutions look to extend the area coverage through LoRa, GSM, and 4G. Others aim to deliver intelligent solutions through ML and CV models [2]. In summary, IoT and AI are widely present technologies in the field and solutions that make use of Neural Networks for disease detection are being developed and proposed. Currently it is already possible to find solutions for disease detection in tomato and corn [7], [8]. Also, it is possible to find solutions using Deep Learning for apple and orange counting [9].

III. System Architecture

This section presents the proposed IoT application for strawberry farming. The IoT application consists of four main components shown in Figure 1.

According to [10], IoT applications in agriculture must fulfill some requirements. First, unpredictable aspects found in agriculture need to be considered in the development phase of agriculture projects. IoT hardware must stay secure against various dangerous environmental aspects in this case. Second, farm size varies from small to large, and IoT solutions need to be scalable and accessible. Third, IoT agriculture applications need to be affordable to every farm that needs to implement the solution. Fourth, IoT agriculture aims to promote sustainability and innovation in the field. Also, the issue of sustainability is important in a global scenario due to economic pressure and fierce global competition.

The development system aims to attain all the requirements cited above. A 3D-printed case is developed to protect the WSN hardware in the agriculture field. The solution is low-cost, precise, and scalable. Also, the application is suitable for large and small farms through LoRa communications. The case material is polylactic acid (PLA), a thermoplastic monomer derived from renewable energy. In addition, the system only uses the essential hardware to capture the critical variables of the crop.

The solution aims to build an intelligent Wireless Sensor Network (WSN), offering farmers a data-driven approach to the strawberry plantation. The applications aim to be suitable for different scenarios and variables, delivering real-time and precise information about the local data through the Internet. Also, an ML algorithm makes real-time classifications of the crop conditions and indicates which crops need irrigation.

The sensor platform offers multiple parallel tasks like temperature, soil moisture, humidity sensing, and weather Internet collecting. Sensors perform the perception function of the IoT strawberry application, capturing various variables from different locations. Data is gathered by sensor nodes spread by the strawberry plantation. Sensors are connected to a microcontroller that controls the intervals of gathering, pre-processing the data of sensors, and transmitting it to the final destination.

The collector node always listens and uploads all the data to the Internet through a Wi-Fi connection and Message Queuing Telemetry Transport (MQTT) protocol. It is a lightweight, efficient, and scalable protocol widely used in IoT applications [11]. The final message stage is on the NodeRED [12], which is a tool for flow-based programming initially developed by IBM to connect hardware devices, Application Programming Interfaces (APIs), and online services as part of the Internet of Things, where all the applications in WSN are connected [13]. All these previous processes are part of the data acquisition pipeline.

WSN and IoT applications make use of multiple devices to collect data at the same time. Also, some agriculture scenarios can require solutions with long-range communication. In this context, LoRa plays a critical role in enabling these applications. LoRa uses spread spectrum communication, and it has a lower probability of collision and interference when compared to similar technologies such as Sigfox. In addition, LoRa does not require a license or signature for use in Brazil, enabling everyone to implement and exploit the service. Futhermore, LoRa is a low-power communication technology. Moreover, the main operating ranges for LoRa are 902-907,5 MHz and 915-928 MHz [14]. All this characteristics increases the scalability and possible scenarios of implementation.

On the other hand, to upload the data to the Internet, a LoRa gateway is needed. Functionalities such as uploading the collected data to the LoRa network are implemented on the collector node. The user can opt for a completely offline network operating with LoRa technology, thus ensuring greater security and privacy of the data available only locally. The LoRa communication settings are transmission frequency of 915 MHz, transmission power of 20 mW, and antenna with a gain of 2.5 dBi. Time Division Multiplexing is a type of multiplexing used in the system, that allows multiple signals to be transmitted in a specific time window within the same physical space, where each signal has its own time and defined bandwidth usage for transmission. Figure 2 illustrates the multiplexing technique used on the application.

The data processing component in collector node is responsible for pre-processing the data for other tasks like database storage, dashboard exhibition, and ML processing. First, the data is stored in the InfluxDB database through a NodeRED Application Programming Interfaces (APIs). InfluxDB is an open-source time-series database used to store the sensor node information. Second, Grafana [18] API gets the InfluxDB data through the API and plots in the cloud and local dashboard. Grafana is an open-source solution that enables flexible and understandable dashboards.

In parallel, ML gets and analyzes important crop metrics such as temperature and humidity. Finally, the ML application uses the Random Forest (RF) algorithm to analyze the sensor and web data for predicting irrigation necessity in the

Fig. 1. Illustration of the proposed system implemented on a strawberry farm.

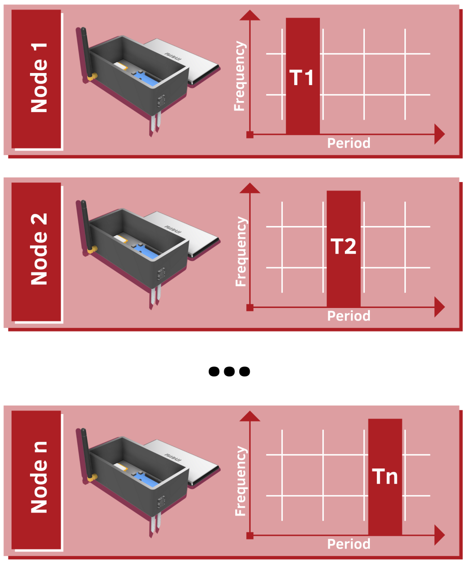

Fig. 2. Multiplexing schema for LoRa communication.

plantation. RFs are a set of tree predictors in which each tree is determined by the values of a random vector sampled separately and uniformly across all trees in the forest [15]. With this approach, the ensemble of n-trees is better than the better tree alone.

The development of the model can be divided into three phases: (I) Data mining, (II) Hyper-parameters setup, and (III+ Training.

The data mining phase refers to finding, cleaning, and organizing the data for the training process. The dataset [16] contains temperature, humidity, soil moisture, wind speed, and other variables. All the features in the dataset not collected by the proposed WSN are removed. Thus, the dataset used for training the model has a reduced size regarding the number of features. After the modifications, the used dataset has 23995 observations and six features: (I) Soil Moisture (%), (II) Temperature (C°), (III) Heat index, (IV) Wind speed (km/h), (V) Air humidity (%), and (VI) Pressure (kPa). Local sensors capture variables related to Soil Moisture, Temperature, and Air humidity. The other variables are captured by API from OpenWeatherMap. The OpenWeather Ltd. owned online service OpenWeatherMap that offers global weather data through API, including real-time weather information, forecasts, projections, and historical weather information for each place [17].

The second phase is related to the model hyperparameter configuration. The RF model is constructed, trained, and exported using the Scikit-Learn library. Some configurations must be made to develop and reach a better and more accurate model. The number of trees generated, the function that measures the split quality, and the trees maximum depth are configured.

Finally, a model is trained and performance metrics captured. A split of the data is done to train and test the model. It is considered the fraction of 70% of the dataset for training, while 30% is used for the test. The RF training does not require much computer power, and because of that, local training is possible.

After training, the model is exported and implemented on NodeRED for real-time data classification. The classifications are inserted into a time-series bar graph, that show to the user how long the strawberry crop was subjected to an irrigation necessity. Data regarding irrigation needs are presented in different colors in a bar format within the dashboard. For example, the green color indicates that the crop does not need irrigation, while the red color indicates the opposite. In this way, the user can better visualize and analyze for which periods the plantation was in water scarcity conditions.

A JavaScript Object Notation (JSON) centralizes all sensor node data and OpenWeatherMap API data into a single package for ML model. This package is organized with columns and values expected by the RF model. The correct organization and formulation of the data are necessary to guarantee ML predictions.

Also, all the collected data can be seen in the dashboard provided on the Grafana platform. A dashboard demo is depicted in Figure 4.

IV. System Validation and Analysis

In this section, some ML and LoRa communication tests are shown in order to validate the proposed system. First, the LoRa tests aim to measure the distance and Received Signal Strength Indication (RSSI) presented in urban and rural areas. Second, the ML tests aim to validate the accuracy of RF model for predicting the most optimal times for irrigation.

The RSSI is the received signal power in milliwatts and is measured in dBm. The measurements are used to verify how well the receiver receives the signal from the sender. A stronger signal stays around -30 dBm, and a weak signal, on

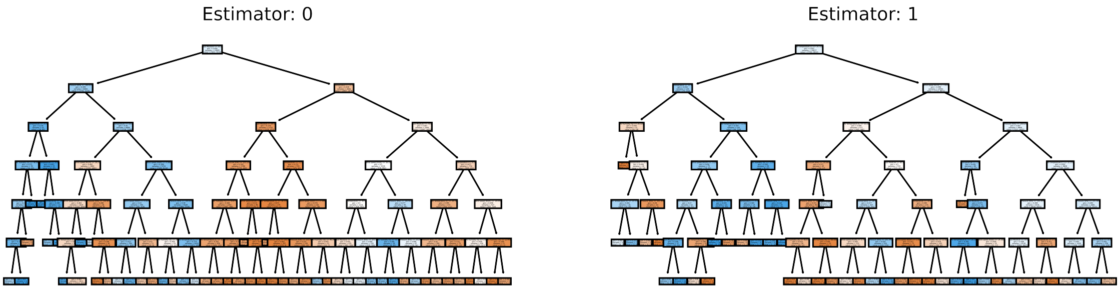

Fig. 3. Ilustration of four trees presents in the Random Forest model.

Fig. 4. Grafana dashboard illustration.

the other hand, stays around -120 dBm. In order to find the maximum distance between the sensor and collector nodes, it established a route with stops every 200 m for measurements. The first test was carried out within an urban perimeter, where houses, trees, and obstacles are present. A rural range test is also conducted to verify its application in both scenarios.

In the test, the Sensor node periodically sends packages to the collector node to check the range and quality of the signal. The package contains the RSSI, temperature, and humidity data. Every 200 m on the route, a record of the RSSI values is made. The route has different reliefs, tree density, and house presence. Because of that, some points may not offer conditions for the receiver to capture the package. Figure 5 shows the planned route for measuring the RSSI.

Fig. 5. The planned route with its respective points for measuring the RSSI indicator.

Table I shows the measured points for the urban scenario and the respective RSSI measures.

During the urban test, the third test point (600m) was critical, due to the shading effects provided by local relief and thus,

TABLE I

RSSI MEASURED IN THE FOUR POINTS.

	Measure 1	Measure 2	Measure 3
Point 1 (200m)	-122dB	-123dB	-121dB
Point 2 (400m)	-122dB	-120dB	-121dB
Point 3 (600m)	–	–	–
Point 4 (800m)	-121dB	-121dB	-122dB

the sensors cannot receive information of this point. Finally, the fourth testing point comprehends the final measuring point, reaching the maximum distance of approximately 800 m. After this milestone, no packet was delivered to the collector node.

The proposed system was also implemented on a real strawberry plantation with a perimeter of 486,04 m and an area of 14,912.4 m^2. LoRa range is tested through four extreme points of the plantation area and respective RSSI annotations. In short, all the areas are covered, but some points present higher RSSI than others. Nevertheless, the signal strength differences do not present problems with delivering the package and information. Figure 6 shows all the reference points of measurement and the fixed spot of the Sensor node. The RSSI measured on all the respective points is shown in Table II.

Fig. 6. Measuring points for LoRa RSSI in strawberry field

TABLE II

RSSI MEASURED IN THEIRS RESPECTIVES POINTS.

	Measure 1	Measure 2	Measure 3
Point 1	-67dB	-67dB	-67dB
Point 2	-75dB	-74dB	-75dB
Point 3	-59dB	-59dB	-60dB
Point 4	-52dB	-51dB	-52dB

The ML model was validated in terms of Area Under the Curve (AUC), which measures and presents the capability of the model to distinguish different classes. The AUC measures the area under the curve of the Receiver Operating Characteristic (ROC). The higher the value presented by the AUC, the greater is the capacity of the model to distinguish between classes. RF presented good metrics related to the AUC, with an AUC metric of 77.1% for the ROC curve presented in Figure 7.

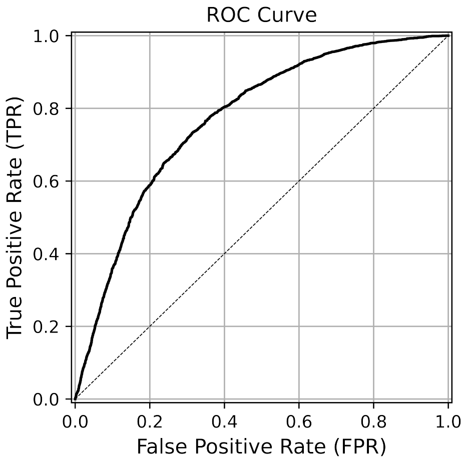

Fig. 7. ROC curve for the proposed ML model.

V. Conclusions

The paper studied and proposed a LoRa WSN with ML capabilities for monitoring strawberry farming. The proposed WSN presents good results in real-strawberry farm and city environments. The ML model presented excellent metrics, enabling to real-time classify and present the best locations and time for irrigation purposes. The overall price of the final system offers possibilities to be implemented on small and large farms. Also, the encapsulation of the hardware uses PLA material, providing minor environmental damage.

The methodology developed to transfer and process data in the strawberry plantation is flexible and can be used, modified, and applied in new scenarios and farms. LoRa communication can be used, others ML models can be employed, other services implemented, different dashboards and database systems used, and much more. The proposed system is scalable, requiring only that the user set a time window for a new Sensor node in the application. On the other hand, the slow nature of the environment variables creates an easy way to do that with numerous sensor nodes. The edge approach of the system also presents possibilities to implement in remote and unconnected scenarios. All the process runs in the collector node, but energy is saved due to LoRa lower-power consumption for transmission and the lower computation power consumption of the Random Forest model.

New machine learning models should be tested in future work. The Scikit Learn library offers other techniques such as K-NN and Support Vector machines for real-time data classification. Adding new sensors to collect more variables in the field should also be considered. This way, it is possible to evaluate new and more accurate methods and models for crop irrigation control. Also, an automatic irrigation control can be implemented within the system. In this way, the RF model would be responsible for analyzing the local variables and activating the irrigation system until irrigation is no longer needed.

References

[1] M. S. Farooq, S. Riaz, A. Abid, K. Abid and M. A. Naeem, "A Survey on the Role of IoT in Agriculture for the Implementation of Smart Farming," *IEEE Access*, vol. 7, pp. 156237-156271, 2019.

[2] M.S. Farooq, S. Riaz, A. Abid, T. Umer, Y.B. Zikria, "Role of IoT Technology in Agriculture: A Systematic Literature Review", *Electronics 2020*, vol. 9, pp. 319, 2020.

[3] S. Devalal and A. Karthikeyan, "LoRa Technology - An Overview," 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2018, pp. 284-290, doi: 10.1109/ICECA.2018.8474715.

[4] C. Cambra, S. Sendra, J. Lloret and L. Garcia, "An IoT service-oriented system for agriculture monitoring," *2017 IEEE International Conference on Communications (ICC)*, pp. 1-6, 2017.

[5] X. Shi, X. An, Q. Zhao, H. Liu, L. Xia, X. Sun, Y. Guo, "State-of-the-Art Internet of Things in Protected Agriculture", *Sensors 2019*, vol 19, pp 1833.

[6] S. Kim, M. Lee, C. Shin, "IoT-Based Strawberry Disease Prediction System for Smart Farming", *Sensors 2018*, vol. 18, pp. 4051, 2018

[7] M. Islam, Anh Dinh, K. Wahid and P. Bhowmik, "Detection of potato diseases using image segmentation and multiclass support vector machine," 2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE), 2017, pp. 1-4, doi: 10.1109/CCECE.2017.7946594.

[8] X. Zhang, Y. Qiao, F. Meng, C. Fan and M. Zhang, "Identification of Maize Leaf Diseases Using Improved Deep Convolutional Neural Networks," in IEEE Access, vol. 6, pp. 30370-30377, 2018, doi: 10.1109/ACCESS.2018.2844405.

[9] S. W. Chen et al., "Counting Apples and Oranges With Deep Learning: A Data-Driven Approach," in IEEE Robotics and Automation Letters, vol. 2, no. 2, pp. 781-788, April 2017, doi: 10.1109/LRA.2017.2651944.

[10] V. S. Konduri, T. J. Vandal, S. Ganguly, A. R. Ganguly, "Data Science for Weather Impacts on Crop Yield", *Frontiers in Sustainable Food Systems*, vol. 4, 2020.

[11] Mqtt.org. 2022. MQTT - The Standard for IoT Messaging. [online] Available at: <https://mqtt.org/> [Accessed 30 May 2022].

[12] G. Tricomi, Z. Benomar, F. Aragona, G. Merlino, F. Longo and A. Puliafito, "A NodeRED-based dashboard to deploy pipelines on top of IoT infrastructure," 2020 IEEE International Conference on Smart Computing (SMARTCOMP), 2020, pp. 122-129, doi: 10.1109/SMARTCOMP50058.2020.00036.

[13] G. Tricomi, Z. Benomar, F. Aragona, G. Merlino, F. Longo and A. Puliafito, "A NodeRED-based dashboard to deploy pipelines on top of IoT infrastructure," 2020 IEEE International Conference on Smart Computing (SMARTCOMP), 2020, pp. 122-129, doi: 10.1109/SMARTCOMP50058.2020.00036.

[14] Agência Nacional de Telecomunicações. 2022. Internet das Coisas. [online] Available at: <https://www.gov.br/anatel/pt-br/regulado/outorga/internet-das-coisas> [Accessed 30 May 2022].

[15] Cutler, A., Cutler, D.R., Stevens, J.R. (2012). Random Forests. In: Zhang, C., Ma, Y. (eds) Ensemble Machine Learning. Springer, Boston, MA. https://doi.org/10.1007/978-1-4419-9326-7_5

[16] Ingle, A. (2020). Crop Recommendation Dataset. Version 1. Retrieved March 20, 2022 from https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset

[17] Openweathermap.org. 2022. urrent weather and forecast - OpenWeatherMap. [online] Available at: <https://openweathermap.org/> [Accessed 28 June 2022].

[18] M. C. Al Fajar and O. N. Samijayani, "Realtime Greenhouse Environment Monitoring Based on LoRaWAnProtocol using Grafana," 2021 International Symposium on Electronics and Smart Devices (ISESD), 2021, pp. 1-5, doi: 10.1109/ISESD53023.2021.9501628.