OBJ OBJ OBJ

Design and implementation of an Embedded Edge-processing Water Quality Monitoring System for underground waters

Christian Correa, *Universidad de Concepción,* Diego Dujovne, *Universidad Diego Portales,* and Fernando Bolaño, *Universidad de Concepción*

Abstract— **Global warming effects are seen around the world and Latin American countries are not an exception, especially for expanding drought areas. Therefore, underground water resources use in the region is incrementing exponentially. However, temporal and spatial underground water information concerning availability and quality is scarce, disabling proper decision-making. In order to close that breach, we propose and Embedded Edge-processing IoT-based Water Quality Monitoring system. This letter introduces the design and implementation of this solution, specifically targeted to monitor irrigation and drinking water extracted from water wells. The system is designed to be deployed in central Chile, considering the topographic conditions, which severely affect power availability and communication resources. The captured data is stored in a data lake, for further processing according to water quality models.**

Index Terms—**Edge processing, IoT, Underground Water, Wireless**

I. INTRODUCTION

Nowadays, in the current scenario of water scarcity and strong climate variability due to global warming, an increasing trend to use underground water has been consolidated as a solution to provide water for both agriculture and human consumption [2],[3]. Nevertheless, this proliferation of new water wells has not been accompanied, at the same speed, by monitoring, management and control systems. Consequently, the Chilean government developed a focused policy to face this problem. The specific objectives of this policy are first, to reduce the spatial resolution of available data, second, to increase the monitoring frequency and third to increase the number of measured variables. In fact, current sampling-based methodologies, which include physical, chemical and biological agent detection processes, provide results with poor spatiotemporal coverage, are labor-intensive and not cost-effective to appropriately detect anomalous events in water composition. In contrast, real-time water quality information enables the authorities to take critical decisions for public health protection [4]. Moreover, an embedded Edge-processing Internet of Things (IoT) solution is capable of providing real-time data, pervasive access, minimum in situ maintenance, low operational cost and scalability to the real-time water quality information systems [10]. However, due to the environmental and topography characteristics present in Chile, Edge processing combined with hybrid WPAN and WLAN network transceivers becomes the right approach to reduce communication traffic and power consumption, which are required to enable connectivity where coverage and energy harvesting resources are limited.

In agreement with the former requirements, this letter introduces the design and implementation of a tailored open-source solution, presenting the hardware and software architecture design. The main architectural systems comprise several water quality sensors, connected to each Embedded IoT node, and a Data Lake.

II. EXISTING SOLUTIONS: BRIEF REVIEW

Existing solutions to water quality monitoring, can be found in [4], [5], [6], [7], [8] and [9]; even commercial solutions are available like Libellium®, Stevens Water Monitoring Systems, Inc and I-real, among others. Authors in [4] describe a typical RaspberryPi data acquisition system, but it has no Edge processing nor a specific event detection algorithm, and uses only 802.11. Sui et al.[5] describe a data capture and transmission system using NB-IoT as the wireless service and no Edge-processing nor a description of a pollution event detection algorithm. NB-IoT as service is not ubiquitous and in general. Kenchannavar et al. [6] briefly describes a water quality measurement system and concentrates on the application of postprocessing ANOVA analysis to detect abnormalities without an Edge processing for near-realtime event detection. Kalamani et al. [7] describe an analysis system based on the definition of water quality classes using fuzzy logic with no Edge processing, nor a related pollution event detection algorithm. Aderemi et al.[8] published a review of groundwater management models based on IoT, defining a common architecture based on a sensor layer, a node layer and an application layer, however, none of them implement Edge processing for pollution event detection. Finally, Wong et al. [9] describe a data capture system with emphasis in the solar power system design, however, the wireless link is based only

Christian Correa is with the Department of Power and Machinery of the Faculty of Agricultural Engineering at the University of Concepción, Av. Vicente Mendez 595, Chillán, Chile (e-mail: ccorrea@udec.cl).
Diego Dujovne, is with the Escuela de Informática y Telecomunicaciones, Universidad Diego Portales, Santiago, Chile (diego.dujovne@mail.udp.cl)
Fernando Bolaño is with the Department of Power and Machinery, the Faculty of Agricultural Engineering, of the University of Concepción, Av. Vicente Mendez 595, Chillán, Chile (e-mail: ferbolano@udec.cl).

in 802.11, and there is no Edge processing and the cloud services are limited to one-year storage and a limited number of graphs. Emerging satellite IoT may be an alternative to achieve connectivity, however, current reliability, cost and power consumption issues make satellite IoT not viable as a medium-term solution, compared to 3G/4G, LoRa/LoRaWAN and WiFi links. Moreover, current IoT satellite constellations provide a long revisit time, which prevents realtime monitoring.

III. WATER QUALITY MODELS

In order to process, storage, and visualize the collected data, we implemented a dedicated platform built as a Data Lake. In particular, this platform allows computing the spatially distributed water quality index, using the standard described for the Canadian Water Quality Index (WQI), which was adopted as standard by the World Health Organization WHO [1].

The WQI equation is calculated using three factors as follows in the (Eq.1):

$$WQI = 100 - \left(\sqrt{F_1{}^2 + F_2{}^2 + F_3{}^2}\right)/1.732. \qquad (1)$$

F₁ represents the percentage of the parameter that exceed the guideline (Eq.2).

$$F_1 = (NFP/TNP).100. \qquad (2)$$

NFP, represents number of failed parameters.

TNP, represents Total number of Parameters.

F_2 represents Frequency: The percentage of individual tests within each parameter that exceeded the guideline (Eq.3).

$$F_2 = (\#failed\ test/Total\ \#\ of\ test).100. \qquad (3)$$

F_3 represents Amplitude: The extend (excursion) to which the failed test exceeds the guideline. In specific, F_3 is calculated in three stages. First, the excursion is calculated (Eq.4).

$$excursion = (FTV/guideline\ value) - 1. \qquad (4)$$

FTV, represents measured field test value.

Second, the normalized sum of excursions (*nse*) is calculated as follows (Eq.5):

$$nse = (\sum excursion\ /Total\ \#\ of\ test). \qquad (5)$$

F_3 is then calculated using a formula that scales the *nse* to range between 1 and 100 (Eq.6):

$$F_3 = nse/(0.01nse + 0.01). \qquad (6)$$

The guideline values are defined for each country. specifically for Chile, the values are stated in the Chilean standard NCh1333 [11]

IV. EVENT DETECTION ALGORITHM

An event is defined as a correlated anomaly in a number of water parameters. These anomalous parameter values are, generally, a result of either natural or human-related actions which lead to a change in water characteristics and composition. The correlation in time among a subgroup of these parameters build a pattern, which is analyzed by the Embedded IoT node. The analysis is based on several time and intensity thresholds, which are defined in either a standard, regulation or guideline. Even though if an event is detected, the alarm is triggered only when a sliding window average of any of the measured parameters exceeds the predefined thresholds. Fig. 1 depicts the flowchart of the process.

First, the incoming value from each type of sensor is filtered by a dual pole digital band pass filter over a sliding window, designed to reduce thermal noise influence and then averaged. The measurement buffer size was estimated to the size of the sliding window plus 10%, so as to absorb possible short-term sensor communication problems. The sliding window and the sampling rate depend on the type of parameter to measure.

Second, after the average is calculated, the result is compared against the reference regulation threshold. If the average is higher than the alarm threshold, the algorithm running on the Embedded IoT Node directly transmits a trigger alert. If not, an event detection algorithm verifies if the parameter value is higher than the regulation threshold. If this is the case, the node starts recording the measured values and implements a maximum value detection by detecting the change in the slope of the signal from positive to negative. When the baseline value is less than the baseline level, the recording of measured values is stopped and the duration and amplitude of the event is transmitted. The parameter characteristics are shown on Fig. 2 for an example-captured signal.

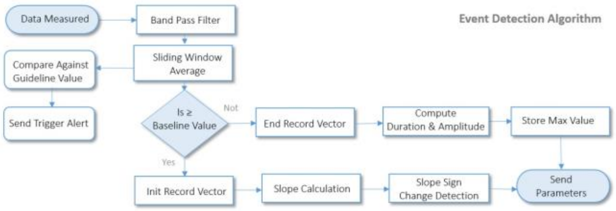

Fig 1. *Block diagram of the Event Detection Algorithm.*

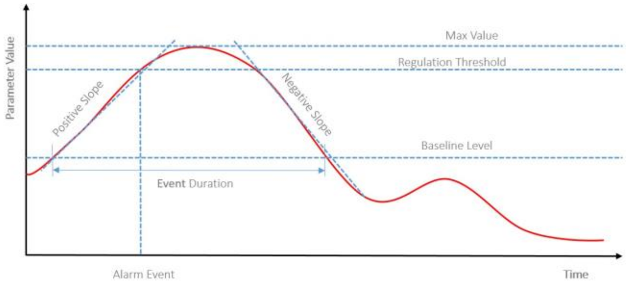

Fig 2. *Parameter feature values calculated at the Edge node.*

With this methodology, the duration, average amplitude and maximum value are transmitted, reducing both transmission power consumption and reducing network resources usage, thus incrementing the node reliability (mostly due to low battery charge due to cloudy days or short winter days) and lifetime.

V. DESIGN AND IMPLEMENTATION

In this section, we provide a description of the functionalities and characteristics of the proposed design. The hardware supporting this system is depicted in Fig. 3, while the system module diagram is described in Fig 4.

First, each **Sensor** captures the selected parameter value. This is scalable among different type of sensors, as described in Fig. 3 on the left: MODBUS sensors can be chained to transmit the values. In this case, MODBUS is run on top of a RS485 bus. Moreover, since sensors are connected using MODBUS, the transmitted packet carries both the measured value and the source sensor ID.

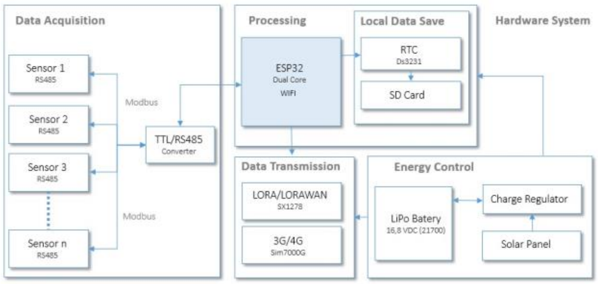

Fig. 3. *IoT sensor node board hardware block diagram.*

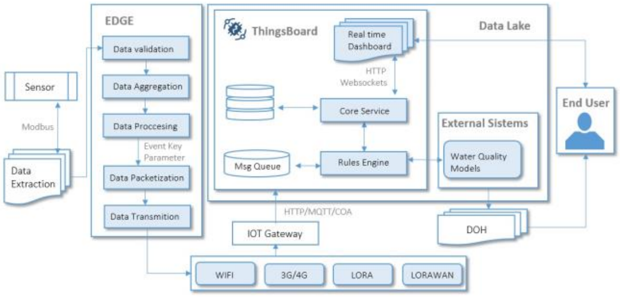

Fig. 4. *Block diagram of the functional modules*

Second, the **Data extraction** process obtains the (Value, ID) pair according to the type of source sensor. Furthermore, different sensors have their own data format, which must be adapted to a single and normalized value. Currently, our proposed design includes the sensors to capture the following parameter values: Chemical Oxygen Demand (COD), Biological Oxygen Demand (BOD), Dissolved Oxygen (DO), pH, Water Depth, Flow, Electric Conductivity (EC) and Total Dissolved Solids (TDS).

The **Edge computing subsystem** comprises a number of preprocessing modules. First, the data extracted from the sensors is validated against the sensor range, to detect possible sensor failure and avoid erroneous data to arrive to the following steps. Second, the Data Aggregation module collects the number of samples required to fill the circular input buffer, for each of the parameters. Third, the data processing module implements an algorithm to detect and process the anomalous water pollution patterns, thus extracting key parameters to be stored in the data lake. Fourth, data is packetized and delivered to the IoT gateway of the embedded IoT node.

Water quality parameters are typically extracted from each individual event, defining the event as a change of magnitude greater than a certain threshold. Consequently, sensors data are acquired, in a fixed frequency, and stored in a timestamped buffer, to detect the events and extract its main features, specifically maximum value, average, duration and frequency. In fact, this strategy was implemented to avoid sending a continuous data flow comprising all parameter values for each sample, but only the input parameters needed to build the water quality models. This strategy reduces the demand for data transmission in one order of magnitude: only a burst of data is transmitted after event detection. A direct consequence of this strategy is reducing both energy consumption and the traffic costs involved in this process. However, several parameters maintain their values for a long time with slight variations, so in this case, their values are sent to the data lake in an hourly basis for long-term analysis.

The **IoT Gateway subsystem** receives the packet from the Edge processing subsystem, and selects the transmission module according to the available technology in the field, building the packet and keeping track of each of the data flows. The transmission modules are enabled only on demand, thus reducing power consumption to the strictly necessary to achieve a successful data transmission process.

Frequently, water wells are placed in harsh environments, in fact, when other kinds of water supply such as public water works are not available, due to topography issues, the typical direct replacement is to deploy a water well. The same topographic conditions generally result in scarce 3G/4G network coverage (mountain ranges and distance from highly populated areas, for example).

The Embedded IoT node boards are based on a dual-core microcontroller ESP 32 Tensilica Xtensa LX6 microcontroller at 240 MHz, from Espressif Systems to provide enough processing power to enable Edge processing and network stack handling, but with low power consumption due to the Idle and Sleep modes implemented. The nodes are also capable of transmitting data using three different technologies: LoRa/LoRaWAN with a 100mW, SX1278 transceiver from Semtech; WiFi (embedded on the ESP32 microcontroller), and 3G/4G with a SIM7500G global-Band LTE-FDD, LTE-TDD, Quad-Band GPRS/EDGE module from Simcom. In order to keep a data backup in case of long-term disconnection or wireless malfunction, we have included internal storage using a microSD card and an RTC for offline timestamping.

The **Power management Subsystem,** as described in Fig. 3, provides stable power for each of the voltage requirements from the processing and communication modules. The battery device is 4S BMS pack for 21700 lithium-ion rechargeable cells and the power source is a 20W solar panel with a DC/DC converter (16.8 VDC to 4.2 VDC) to power the microcontroller board. Topographic conditions also limit the real power availability due to implementation limitations. For example, the solar panels must be selected with a limited size and weight, and sunlight can be even restricted to a few hours per day due to mountain ranges and trees, even in summer. Moreover, our experience shows that a robust and reliable power management system is required due to the low-quality energy characteristics in rural environments, such as wide voltage variations (AC), intermittent sources and high electromagnetic interference, due to the motor pumps and other actuators sharing the same power line.

The **Data Lake** at the server side was implemented on top of the Community version of Thingsboard, an Open-source IoT platform which provides device management, data collection, processing and visualization for IoT solutions. We analyze the incoming telemetry data and trigger alarms with a customized event processing design. Control devices use remote procedure calls (RPC) and workflows are based on a device life-cycle event, REST API events and RPC requests.

This platform was implemented to act as a Data Lake in order to provide, through an API, source data for a tailored platform developed to implement the water quality models and to send the analytic results to the Water Works Department (DOH, in its Spanish acronym), the governmental office, advocated to maintain water quality and availability records.

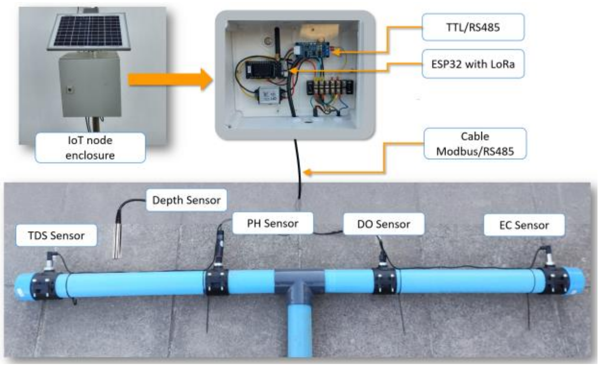

Fig 5. *Sensor module implementation and its wire diagram connection to the Embedded IoT node*

Finally, Fig. 5 shows the prototype implementation in detail. The Embedded IoT node electronic components are encased in an IP67 box, where the most relevant hardware modules can be recognized. The water characterization sensors are installed along a tube where the underground water flows, reaching each of the sensors with the same speed and pressure, so PH, Dissolved Oxygen, Total Dissolved Solids and Electrical Conductivity sensors share the same measurement conditions. The Depth and temperature sensors are located inside the water well, since these features are directly related to the water source.

Fig 6. *A controlled experiment to test the system behavior.*

In order to illustrate the behavior of the algorithm, we performed a controlled experiment by adding a pollutant to the water in order to change the electroconductivity and generate an event, which is detected at the timestamp 19:56:38. Fig.6 shows the Electroconductivity value during the test event, the calculated slope and the event itself. We used a Sliding Window of 30 for this parameter.

VI. CONCLUSION

We have described the design and implementation of an Embedded Edge-processing Water Quality Monitoring system, where the design criteria respond to traffic, reliability and power consumption requirements to operate in underground water wells. The implementation complies with international water quality standards, by obtaining key parameter values at the edge nodes, which are transmitted to a Data Lake. Data transmission is achieved by including four different wireless modules in an IoT Gateway subsystem, so as to provide flexibility for different kinds of configurations and take advantage of local networks. The implementation is based on a hardware board, which includes also power management, battery storage and an external solar panel.

ACKNOWLEDGMENTS

This work was supported in part by Corfo, Water Technology Consortium CoTH20. 20CTECGH-145896, CRHIAM Conicyt/Fondap/15130015 project, the Conicyt Advanced Human Capital Program of the Government of Chile, by the CI4 Center for Industry 4.0, CYTED AgIoT Project (520rt0011) and Proyecto Asociativo UDP "Plataformas Digitales como modelo organizacional".

REFERENCES

[1] C.J. Rickwood, G.M. Carr, Development and sensitivity analysis of a global drinking water quality index. Environ Monit Assess 156, 73 (2009).

[2] G. Donoso, E. Lictevout, J.D. Rinaudo, 2020. Groundwater Management Lessons from Chile. Global Issues in Water Policy, 24:481-509.

[3] J.D. Rinaudo, and G. Donoso, 2019. State, market or community failure? Untangling the determinants of groundwater depletion in Copiapó (Chile). Int. J. of Water Resources Development, 35(2), pp.283-304.

[4] N. Vijayakumar and R. Ramya, "The real time monitoring of water quality in IoT environment," 2015 Int. Conf. on Innovations in Information, Embedded and Communication Systems (ICIIECS), 2015, pp. 1-5,

[5] H. Sui, G. Zheng, J. Zhou, H. Li et al. "Application of NB-IoT Technology in City Open Water Monitoring," 2020 6th Int. Symp. on System and Software Reliability (ISSSR), 2020, pp. 95-98,

[6] H. H. Kenchannavar, P. M. Pujar, R. M. Kulkarni and U. P. Kulkarni, "Evaluation and Analysis of Goodness of Fit for Water Quality Parameters using Linear Regression through the Internet of Things (IoT) based Water Quality Monitoring System," in IEEE Internet of Things J.,

[7] P. Kalamani and N. Rakesh, "Analysis of Lake Water Quality classification using Fuzzy inference Systems," 2018 2nd Int. Conf. on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2018 2nd Int. Conf. on, 2018, pp. 168-173,

[8] B.A Aderemi, Olwal, T.O.; Ndambuki, J.M.; Rwanga, S.S. A Review of Groundwater Management Models with a Focus on IoT-Based Systems. Sustainability 2022, 14, 148.

[9] J. W. Yong, N. Rei, S. Yoshihisa K. Akinori, et al., Toward industrial revolution 4.0: Development, validation, and application of 3D-printed IoT-based water quality monitoring system, J. of Cleaner Production, vol. 324, 2021, 129230, ISSN 0959-6526, https://doi.org/10.1016/j.jclepro.2021.129230.

[10] P. Velásquez, L. Vásquez, C. Correa and D. Rivera, "A low-cost IoT based environmental monitoring system. A citizen approach to pollution awareness," 2017 Chilean Conf. on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON), 2017, pp. 1-6

[11] Water quality requirements for different uses NCh1333:1978 Mod.1987: Instituto Nacional de Normalización Chile. INN, May 1987.