

Real time monitoring of water Quality using IoT and Deep learning

Saloua SENHAJI^{1,2}, Mohamed HAMLICH² and Mohammed OUZZANI JAMIL¹

¹Laboratoire Systèmes et Environnement Durables (SED), Faculté des sciences de l'ingénieur (FSI) Université Privée de Fès (UPF), Fez, Morocco

²Laboratoire Ingénierie Des Structures, Systèmes Intelligents Et Energie Electrique, ENSAM ,Université Hassan II, Casablanca, Morocco

salouasenhaji@gmail.com

Abstract. Access to safe drinking water is one of the most pressing issues facing many developing countries. Water must meet Environmental Protection Agency (E.P.A.) requirements. The normal method of measuring physico-chemical parameters is to take samples manually and send them to the laboratory to check the water quality. In this paper, we proposed a new intelligent design of a real-time water quality monitoring system using Deep Learning technology. This system is composed of several sensors that allow us to measure water parameters (physico-chemical parameters), bacteriological parameters and organoleptic parameters) and to detect the presence of certain substances (undesirable substances, toxic substances) and of a single-board/mobile computer module, Internet and other accessories. Water parameters are automatically detected by the single-board computer. Raspberry Pi3 model B. The single board computer receives the data from the sensors and this data is sent to the web server using the Internet module. It is able to detect the water quality situation worldwide. The data will be analysed in real time. The application of deep learning to these areas has been an important research topic. The Long-Short Term Memory (LSTM) network has been shown to be well suited for processing and predicting large events with long intervals and delays in the time series. LSTM networks have the ability to retain long-term memory.

Keyword: water quality, real time, Deep Learning, IoT, Raspberry.

1. Introduction:

The surface water resources used for drinking water production in Morocco are different in nature depending on their origin and the anthropic impacts they receive. Thus, drinking water producers aim to provide consumers with water that will not pose a risk to human health and that will be compatible with Moroccan food water quality standards [1, 2]. Thus, these standards impose quality requirements on surface water used for drinking water production for a given level of treatment. New environmental policies require improved methods for the study and assessment of well water quality, in this context we propose an intelligent well monitoring technique to automate and computerise water policing tasks.

In this context, this manuscript presents a new design of a real-time water quality monitoring system. The system is composed of several sensors that will be used to measure chemical parameters of water such as organoleptic parameters, physico-chemical parameters, undesirable substances, toxic substances. The values

measured by the sensors [6] will be processed by a controller. The system will generate large amounts of data that we will be forced to process in real time, to solve this problem we propose deep learning techniques. Deep learning is a popular approach to machine learning that has made many advances in all traditional areas of machine learning [8, 10, 11, 12].

The Internet of Things (IoT) [3, 5, 7] and smart environment deployments generate large amounts of time-series sensor data that need to be analysed. The application of deep learning to these areas has been an important research topic. The Long-Short Term Memory (LSTM) network has proven to be well suited to process and predict large events with long intervals and time series delays [4]. LSTM networks have the ability to maintain a long-term memory. In an LSTM network, a stacked LSTM hidden layer also allows learning a high-level temporal feature without the need for a hidden layer and without the need for fine-tuning and pre-processing that would be required by other techniques.

2. PROBLEM DEFINITION

Environmental regulations require the monitoring of the environmental status of well water, in order to monitor water quality.

The current system collects water parameters in a manual way, using specialised sensors, the collected data is sent to the laboratory. The process of moving and collecting data takes hours, resulting in a delay in acting on sources of poor water quality, thus poor accuracy of the current status of the wells. A large number of parameters are tested in the current system, such as pH, turbidity, conductivity, salinity, phosphorus and nitrogen.

The shortcomings of the existing system : long and less efficient, high cost, more labour intensive

3. PROPOSED METHODOLOGY

This section presents the overall block diagram of the of the proposed system. The block diagram is shown in Figure 1. This block diagram includes a good number of devices with specific sensors, and the data collected from all the devices is collected and sent to the Raspberry Pi 3 Model B.

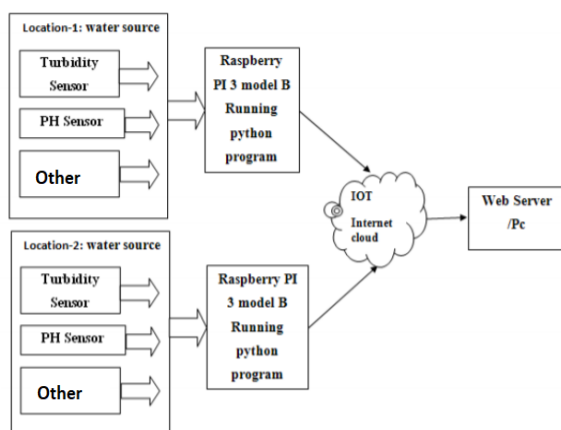


Fig 1. Overall diagram of the proposed system

The device consists of several sensors to measure water quality parameters. The data from the sensors is sent directly to the Raspberry pi3 model B. Thus, the proposed system receives the data from the sensors and processes it, puts the data in a text file which is transmitted to the IOT.

To transmit the data to the IOT, a gateway is created on the Raspberry pi 3 model B using file transfer protocol (FTP). In the proposed system, to monitor the processed data over the internet, cloud computing technology is used which provides the personal local server. In cloud computing a separate IP address is provided, which allows monitoring the data from anywhere in the world using the Internet. and to analyze the data we have Deep Learning techniques more precisely LSTM (Long-Short Term Memory) . To access this data and make the system user friendly, a browser application is provided which runs on HTTP so the user can access and monitor the data from anywhere in the world.

4. Water quality

In this document, surface water is defined as untreated or unfiltered water from wells that water utilities or individuals use for drinking. Treated water is water that is delivered to consumers after being purified. In general, the minimum treatment includes disinfection. Quality drinking water is ultimately defined as that which is safe for drinking and cooking (Gadgil 1998). The subjectivity associated with such a holistic definition has led to the functional separation of water quality into Five categories: (1) water free of pathogenic organisms (2) water containing harmful chemicals below defined thresholds and physical parameters within acceptable limits, and (3) water containing radioactive compounds below defined thresholds (Health Canada 1996).

Figure 2 gives a summary of the classification grid of waters used for the production of drinking water.

category	A1		A2		A3	
	G	I	G	I	G	I
organoleptic parameters						
colors	mg pt/l	<10	20	50	100	200
odour at 25 C		<3	-	10	-	20
physico-chemical parameters						
Temperature	°C	20	30	20	30	20
PH	pH	6,5-8,5	-	6,5-9,2	-	6,5-9,2
conductivity at 20 degre C		1300	2700	1300	2700	1300
chlorides	mg/l	300	750	300	750	300
Sulphates	mg/l	200	-	200	-	200
suspended solids	mg/l	50	-	1000	-	2000
Dissolved O2	mg/l	7 (90%)	-	5 (70%)	-	3 (50%)
DBO5	mg/l	3	-	7	-	10
DCO	mg/l	-	-	25	-	40
oxidability	mg/l	2	-	5	-	10
undesirable substances						
bore	mg/l	-	1	-	1	-
ammonium	mg/l	0,05	0,5	1	1,5	2
NTK	mg/l	1	-	2	-	3
nitrates	mg/l	-	5	-	50	-
phosphore	mg/l	0,4	-	0,7	-	0,7
baryum	mg/l	-	1	-	1	-
cuivre	mg/l	-	1	-	2	-
zinc	mg/l	-	50	-	5	-
manganèse	mg/l	-	0,1	0,1	0,1	1
fer dissous	mg/l	-	0,3	1	2	1
fluorures	mg/l	0,7	1,5	0,7	1,5	0,7
hydrocarbures dissous	mg/l	-	0,05	-	0,2	0,5
phenols	mg/l	-	0,01	-	0,05	-
detergents anioniques	mg/l	-	0,5	-	0,5	-
toxic substance						
arsenic	µg/l	-	50	-	50	-
cadmium	µg/l	1	5	1	5	-
chrome total	µg/l	-	50	-	50	-
plomb	µg/l	-	50	-	50	-
mercure	µg/l	-	1	-	1	-
selenium	µg/l	-	10	-	10	-
nickel	µg/l	-	50	-	50	-
cyanures	µg/l	-	50	-	50	-
pesticides substance	µg/l	-	0,1	-	0,1	-
pesticides au total	µg/l	-	0,5	-	0,5	-
HPA	µg/l	-	0,2	-	0,2	-
bacteriological parameters						
coliformes fecaux	/100ml	20	-	2000	-	2000
coliformes totaux	/100ml	50	-	5000	-	5000
streptocoques fecaux	/100ml	20	-	1000	-	1000

Category A1 for water requiring simple physical treatment and disinfection, in particular by filtration and disinfection, to be drinkable.

Category A2 for water requiring, in order to be drinkable, normal physical and chemical treatment and disinfection, in particular by prechlorination, coagulation, flocculation, decantation, filtration and disinfection (final chlorination) decantation, filtration and disinfection (final chlorination).

Category A3 for water requiring, in order to be drinkable, advanced physical and chemical treatment, refining and disinfection, in particular by "break-point" chlorination, coagulation, flocculation, decantation, filtration, refining (activated carbon), and disinfection (ozone, final chlorination).

Up to now, the treatments are done in a traditional way with a displacement of specialists, water police, where a loss of time, a delay of decision making, and additional expenses.

To remedy these problems, we propose an intelligent solution to collect and analyse data in real time from various sensors, thus allowing preventive maintenance of the water purification system.

5 . The Internet of Things

The Internet of Things (IoT) [7, 8, 9] has been identified as one of the key developments in the technology portfolio. The Internet of Things (IOT) enables the interconnection of communicating objects that are placed in various locations and may be distant from each other. The Internet of Things is a concept in which network devices have the ability to collect and sense data in the world and then exchange it over the Internet where it is exploited and processed for various purposes. The Internet of Things represents a vision in which objects are an integral part of the Internet: each object is uniquely identified and has access to the network. The IoT is very different from traditional human-to-human communication, which represents a considerable challenge for existing telecommunications and infrastructure. In addition, the IoT provides immediate information on access to physical objects with great efficiency. The concept of the Internet of Things is very useful for real-time monitoring of sensor data. The Internet of Things (IoT) is a kind of network technology, which is based on information sensing equipment such as infrared sensors, GPS, etc., where all objects join the Internet to exchange information for intelligent identification, location and tracking.

In the proposed system, we introduce a cloud computing technique to monitor sensor values over the Internet. Cloud computing allows applications to be accessed as utilities over the Internet. The characteristic site of cloud computing and its development are explained. Cloud computing is a large-scale processing unit that processes in real time and is also a very low cost IP technology.

6. Deep Learning for Anomaly Detection in IoT Data

6.1 The difficulties of detecting anomalies in time series

Time series measurements have a high dimension, difficulty, sensitivity and disturbance level. If data mining is carried out on the basis of the original time series, not only will the storage and computational costs be significant, but the relevance and accuracy of the algorithm will also be affected. The question is how to prepare the time series data efficiently, avoiding the destruction of key information in the data. Limiting the dimensionality of the data and reducing noise are the main objectives of preprocessing. Noisy data increase the complexity of an anomaly detection system on given time series. In addition, when abnormal data are not present or are few in number, it is very difficult to determine the classification pattern of normal and abnormal sequences.

6.2 The LSTM-Gauss-NBayes model to detect anomalies

To address the challenges of processing time series data, we first use the technique of downsampling to obtain the characteristic subsequence of the original time series. Top-down sampling reduces the number of dimensions of the original time series and facilitates model learning. At the same time, in order to accelerate the convergence rate of the model, we normalise our data. the model using the min-max normalisation for time series data, which is a linear transformation of the original data. The transformed values are mapped onto the interval [0, 1].

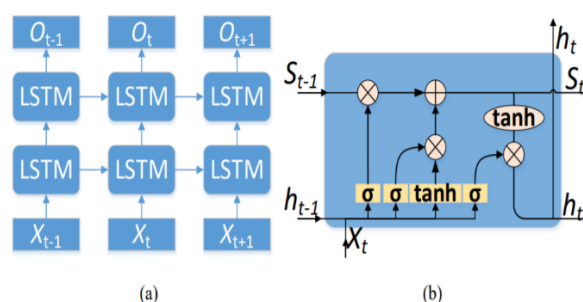


Fig 3. LSTM prediction model

The Figure 3 (a) illustrates the expansion of the LSTM prediction model, where the hidden layer LSTM units are fully connected by recurrent connections. The LSTM units in the hidden layer are fully connected by recurrent connections. Each module of the bottom LSTM hidden layer in the overlay LSTM layer is connected to each cell of the LSTM hidden layer above it through recurrent connections.

The bottom LSTM cell in the overlaying LSTM layer is connected to each cell in the hidden LSTM layer above it via a direct connection. Furthermore, Figure 3 (b) shows the internal structure of the LSTM layer, where σ and \tanh represent the activation function.

Conclusion

In this paper, we propose a combination of long-term and short-term memory IoT neural networks for real time water monitoring. Our proposed system can monitor water quality automatically, and it is updated on the website servers at low cost and does not require staff in service. Water quality testing should therefore be low cost, efficient and fast. Operation is simple, the system can be extended to monitor hydrology, air pollution, industrial and agricultural production, etc. It has wide application and extension value.

REFERENCES

1. Mo Deqing, Zhao Ying, Chen Shangsong, "Automatic Measurement and Reporting System of Water Quality Based on GSM," 2012 International Conference on Intelligent System Design and Engineering Application (2012).
2. Mijović S. and Palmar B. (2009). "Continuous water quality monitoring on river Kolubara" ,30th scientific professional meeting with international participation.
3. George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel, Greta M. Ljung., "Time series analysis: forecasting and control (Fifth Edition)," John Wiley & Sons, 2015.
4. D. Wu, Z. Jiang, X. Xie, X. Wei, W. Yu and R. Li, "LSTM Learning With Bayesian and Gaussian Processing for Anomaly Detection in Industrial IoT," in IEEE Transactions on Industrial Informatics, vol. 16, no. 8, pp. 5244-5253, Aug. 2020.
5. Z. Qin, D. Wu, Z. Xiao, B. Fu and Z. Qin, "Modeling and Analysis of Data Aggregation From Convergecast in Mobile Sensor Networks for Industrial IoT," in IEEE Transactions on Industrial Informatics, vol. 14, no. 10, pp. 4457-4467, Oct. 2018.
6. Sachin S.patil, S.J.Patil, M.M.Raste "Air Pollutant Monitoring Using Sensor Networks" International Journal Of Advanced Research in Electronics and Communication Engineering, Volume.3.Issue.8.,Aug-2014,pp..829..-.833.
7. S.Senhaji, M.Hamlich, IoT technologies and Deep Learning to support the smart building development: Review, Opportunities, and Challenges, International Network of Research and Training, INRT (2021).
8. S. Senhaji, Sanaa Faquir and Mohammed Ouazzani Jamil. Towards robotics and artificial intelligence for the prevention of Covid 19 pandemic, E3S Web of conferences, 229, 01035. (2021)
9. S.Senhaji ,S.Faquir, F. Harchli, T. Hajji and M. Ouazzani Jamil, A new method to Analysis of Internet of Things Malware using Image Texture component and Machine Learning Techniques, The First Edition of the International Conference on Artificial Intelligence & Industrial Applications (2020) .
10. Hamlich 1.,M Ramdani, "Improved ant colony algorithms for data classification", Proceedings of 2012 International Conference on Complex Systems, ICCS 2012, 2012, 6458563.
11. M. Hamlich 2, M. Ramdani, « Fuzzy Ant Miner», Proceedings of the IADIS International Conference Intelligent Systems and Agents 2012, ISA 2012, IADIS European Conference on Data Mining 2012, ECDM 2012, 2012, pp. 142–149.
12. M. Hamlich, A. El Khantach, Belbounaguia, N., « Machine learning methods against false data injection in smart grid», International Journal of Reasoning-based Intelligent Systems, 2020, 12(1), pp. 51–59