Water Quality Monitoring System using IoT and Machine Learning

Nikhil Kumar Koditala
Computer Science and Engineering
BML Munjal University
Gurugram, India
nikhil.koditala.15cse@bml.edu.in

Dr.Purnendu Shekar Pandey
Computer Science and Engineering
BML Munjal University
Gurugram, India
purnendu.pandey@bml.edu.in

***Abstract*—World Economic Forum ranked drinking water crisis as one of the global risk, due to which around 200 children are dying per day. Drinking unsafe water alone causes around 3.4 million deaths per year. Despite the advancements in technology, sufficient quality measures are not present to measure the quality of drinking water. By focusing on the above issue, this paper proposes a low cost water quality monitoring system using emerging technologies such as IoT, Machine Learning and Cloud Computing which can replace traditional way of quality monitoring. This helps in saving people of rural areas from various dangerous diseases such as fluorosis, bone deformities etc. The proposed model also has a capacity to control temperature of water and adjusts it so as to suit environment temperature. Based on our model we have achieved R-squared score of 0.933.**

***Index Terms*—Water quality monitoring, Internet of Things, Machine learning, Cloud Computing, MQTT, Rural Development**

I. Introduction

Nowadays, Due to limited water resources and increasing population water has become a vital resource for mankind. Clean and safe drinking water is the most important resource for mankind. As most of the diseases these days spread through water there is a need for online real time water quality monitoring system. The methods used for water quality assessment at present involve collection of random samples of water at various locations weekly or monthly and analyzing them in the laboratories. This approach is not much efficient because they have various drawbacks such as long time consumption, only water samples from few areas can be determined simultaneously. This method also involves manual work to monitor the quality of water regularly. Theses methods are also costly and are not capable enough in large populated countries like India, China. In order to overcome these drawbacks we need a real time system which monitors water quality through sensors such as pH, turbidity and temperature and updates those values in Cloud service. This system consists of sensors which measure the chemical composition of water. These sensor values are then passed to NodeMCU micro controller which has inbuilt WiFi module, using which the data is passed over to Azure Event Hub. From Event Hub data is stored in Azure Storage hub in the form of structured data.

Thereafter using Stream hub data is streamed to external services. PowerBI which is also a Microsoft platform is used to display the sensor values in the form of Web page. This paper also uses MQTT client broker architecture to transmit data from micro controller to external MQTT broker service. Turbidity measures the large number of suspended particles in water that is invisible. Higher the turbidity higher the risk of diarrhea, cholera. Lower the turbidity then the water is clean. besides turbidity, pH is also an important measure which measures the acidic level of drinking water. Temperature sensor measures how the water is, hot or cold.

Other part of this paper is to sense the external temperature near the water storage and control heater or cooler respectively depending on temperature. This part of paper uses machine learning, where the system predicts the weather conditions using previous labeled dataset and controls heater and cooler according to external weather conditions. This makes the system completely automated without any manual interventions. Whenever the value of turbidity reaches predefined threshold an email alert will be sent to concerned authorities informing the situation, forcing them to take immediate action.

II. Literature Review

Pradeepkumar M, Monisha J, Pravenisha R, Praiselin V, Suganya Devi K entitled ”The Real Time Monitoring of Water Quality in IoT Environment”. This paper discusses not only sensor based system but also it introduces cloud computing architecture into IoT which makes the sensor data accessible worldwide.

Atif A, Wasai Shadab, Mohammad Hassan, Shamim, Alelaiwi and Anwar Hossain entitled ”A Survey on Sensor-Cloud: Architecture, Applications, and Approaches” discusses about the sensor-cloud infrastructure, approaches, and different layers of transferring generated data by connecting sensors with cloud services.

Nikhil Kedia entitled ”Water Quality Monitoring for Rural Areas-A Sensor Cloud Based Economical Project” This paper not only highlights embedded sensor systems, but also discusses the challenges and economic viability of the system involving Mobile Network Operator and Government. This system directly contacts Government to take action based on the severity of quality issue.

R.Karthik Kumar, M.Chandra Mohan, S.Vengateshapandiyan,

978-1-5386-2599-6/18/$31.00 ©2018 IEEE

M.Mathan Kumar, R.Eswaran entitled âSolar based advanced water quality monitoring system using wireless sensor networkâ uses solar node to power the wireless sensor network and displays results using GUI created through Matlab.
Jayti bhatt, Jignesh patoliya entitled ”IoT based water quality monitoring system”. This paper shows the design of water quality monitoring system using pH, turbidity, dissolved oxygen and temperature sensor. It uses raspberry Pi as core controller and used Zigbee protocol to transmit data remotely.
Fiona Regan, Antoin Lawlor and Audrey McCarthy, ”Smart Coast Projectâ Smart Water Quality Monitoring System”, designed smart water quality monitoring system. In that system they made water quality smart sensors so the sensors send data wireless to the device which collects data from all the nodes. This system is highly scalable, faster and user friendly, but it is costly because of type of sensors used.
Vaishnavi V. Daigavane, Dr. M.A Gaikwad entitled ”Water Quality Monitoring System Based on IOT”. This paper measures Turbidity, pH and also flow of water using flow sensor. This paper shows the most economical and convenient method of water monitoring system by using the existing GSM network to transmit sensor values.
ZulhaniRasin and Mohd Rizal Abdullah, ”Water Quality Monitoring System Using Zigbee Based Wireless Sensor Network”. This paper uses ZigBee based wireless sensor network to develop water quality monitoring system. This paper also shows the usage of c++ to create a GUI to make the data publicly accessible.

III. IoT Architecture

This section presents the IoT architecture based on which the system is built. Fig. 1 explains the layer wise architecture of our model according to reference model of IoT world Forum [2] from layer 1 at bottom to layer 7 at top.

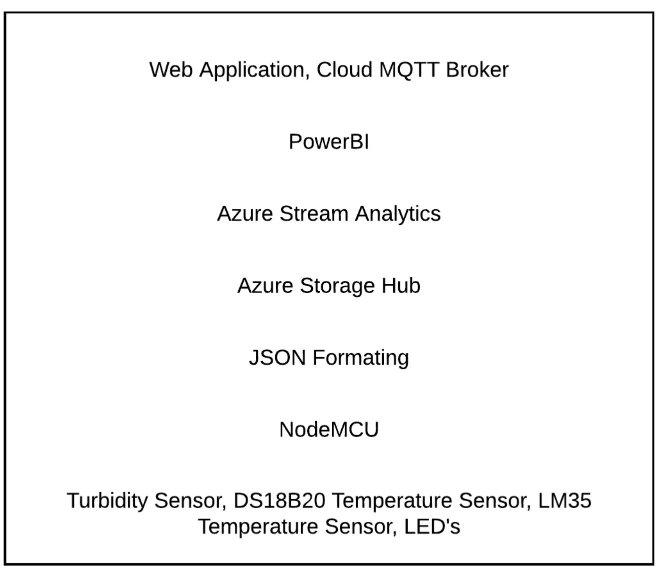

Fig. 1. IoT Architecture

- Layer 1: Layer 1 consists of all the sensors such as Turbidity sensor, Temperature sensor and actuators such as heater or cooler LEDs.
- Layer 2: To ensure reliable communication between devices and cloud we choose WiFi as connecting technology such that there is reliable delivery of messages across the network.
- Layer 3: Azure Event hub only accepts the data which is in structured format. So, Before sending the data to cloud, it is formatted to a JSON packet and then sent to Azure Event Hub.
- Layer 4: For data accumulation we are using Azure storage hub to store the data on cloud servers which ensures reliability and cost effective. Several instances of data is also made which ensures backup ability in case of physical damage of main server.
- Layer 5: Data from Storage hub is transmitted to Azure stream analytic which makes data accessible to other devices outside azure who have access to data.
- Layer 6: For analysis and reporting we are using PowerBI which is a Microsoft tool for data analysis and visualization part of data.
- Layer 7: We built a web application where user can enter his credentials and login into web page, where he can access all the visual charts which depict the quality of water. Besides web page to demonstrate working of MQTT we used an MQTT broker to transmit the data from client which is our gateway to broker which is cloudMQTT.

IV. Proposed Model

This section explains the algorithms used in our system for adjusting temperature and e-mail sending algorithm.

A. Adjusting Temperature

Algorithm. 1 demonstrates the temperature control module of the system. External temperature is calculated from getExternal() function, which is explained in detail at machine learning section. We also calculate the temperature of water using DS18B20 waterproof temperature sensor. If the external temperature is hot i.e. greater than $25\,^{\circ}\mathrm{C}$ and water temperature is hot($>15\,^{\circ}\mathrm{C}$) then cooler gets turned ON and heater gets turned OFF. Similarly in cold climates the second if statement controls the temperature.

B. Sending Email

Algorithm. 2 explains the email module of the system. Turbidity is calculated from getTurbidity() function and whenever the turbidity is greater than 4.0, the system sends the email to

Result: Water temperature adjusted according to Environment
External temperature = getExternal();
Temperature = getTemperature();
if $Externaltemperature > 25$ **then**
  if $temperature > 15$ **then**
    heater = OFF;
    cooler = ON;
  end
end
if $Externaltemperature < 10$ **then**
  if $temperature < 15$ **then**
    heater = ON;
    cooler = OFF;
  end
end

Algorithm 1: Algorithm for adjusting temperature

concerned authorities. At the same time it also keeps a check on count so that the authorities are not flooded with emails continuously.

Result: Sending email to authorities
turbidity = getTurbidity();
count = Threshold limit;
if $turbidity > 4.0$ **then**
  if $count > 0$ **then**
    sendEmail();
    count = count - 1;
  end
  polluted = true;
end

Algorithm 2: Algorithm for sending Email to authorities

V. Machine Learning

Providing temperature controlled water is also important besides quality, in areas near the equator due to extreme climatic conditions. This model uses machine learning to predict temperature of environment near the water storage system.

A. Data Source

The dataset [8] contains monthly temperature records across all major cities all over the world from Jan 1701 to May 2011. The dataset is scrapped from DataMarket. Table I shows the top 5 rows of the dataset. The first coloumn refers to year, second refers to month and the third refers to Temperature at the specific city.

B. Algorithm

In multi variable linear regression the target value is expected to be a linear combination of the input variables.
In mathematical notation, if is the predicted value.Across the

TABLE I
Monthly Average of New Delhi, India

Year	Month	New Delhi
1931	1	15.4
1931	2	15.2
1931	3	21.2
1931	4	29.7
1931	5	33.1

module, we designate the vector as $coef$ and as intercept.

$$y(w,x) = w^0x^0 + w^1x^1 + ...w^{\mathrm{n}}x^{\mathrm{n}}$$

According to our dataset, sklearn linear regression module predicted $coef$ as [-0.00056244 , 0.19589288] and intercept as 24.80431407431011. So, our equation to predict temperature is as follows:
y(year, month) = 24.80431407431011 + year(-0.00056244) + month(0.19589288)

VI. Experiment And Simulation Setup

As this paper involves various technologies such as IoT, machine learning and cloud computing it requires a large hardware and software pipe lining. This section explains the hardware and software setup of the model.

A. Block Diagram

Fig 2 explains the complete structure of the project in the form of a block diagram. The block diagram consists of temperature sensor, turbidity sensor, external temperature sensor, NodeMCU microcontroller, power supply, Azure cloud platform and a webapp built through powerBI. Azure cloud platform has 3 sub blocks, Event hub, Storage unit and Stream hub. The event hub acts as a first line of interface between the sensors and azure cloud. Event hub then transfers the data to storage unit of azure. Then we use stream hub to export the data to powerBI.

B. Hardware and Software Components

Different hardware and software components used in this paper are listed below:

1) NodeMCU: NodeMCU is a microcontroller with inbuilt wifi module present. This makes it suitable for IoT models involving less number of sensors or actuators. it runs on 3.3v. It has 17 GPIO pins and 1 analog pin. It supports 802.11 b/g/n. It is also economical when compared to other microcontrollers.

2) Turbidity Sensor: Turbidity is the quantitative measure of suspended particles in a fluid. When the value of turbidity is less than 3.5 NTU then water is said to have dissolved solids in it. If it is greater than 3.5 NTU then water is safe to drink. It consumes operating current of 40mA.

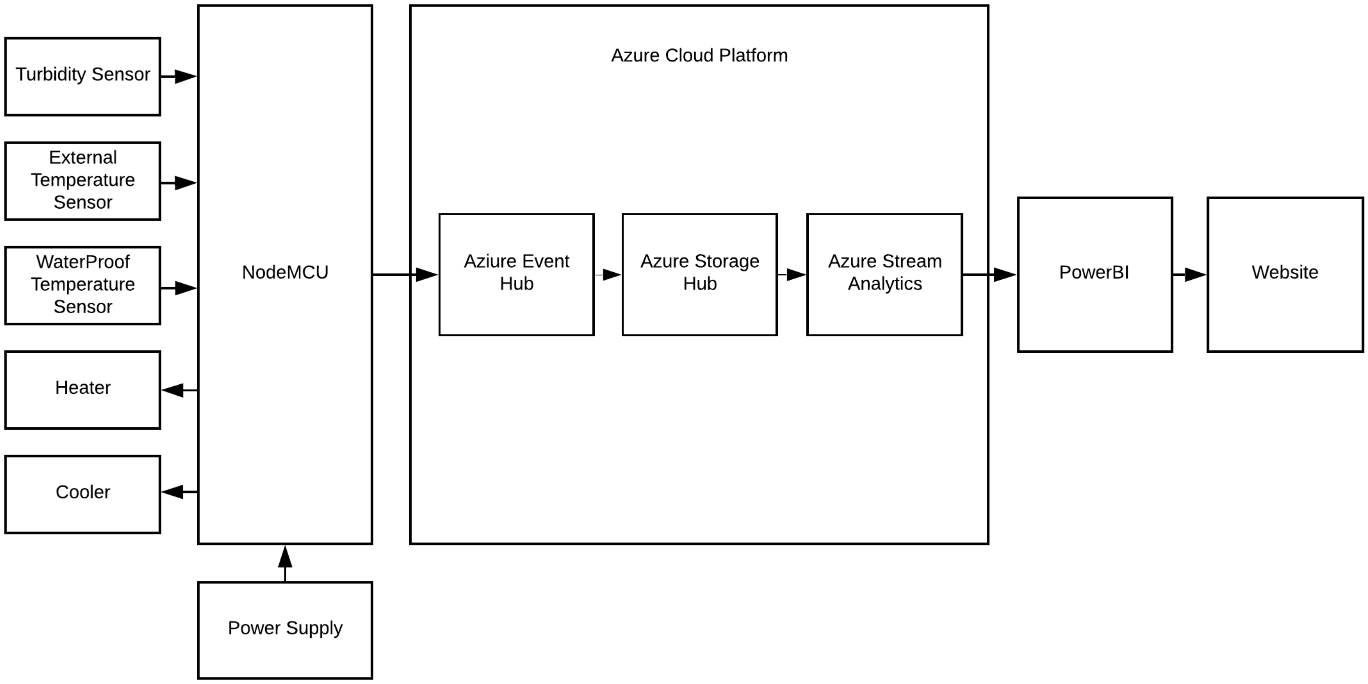

Fig. 2. Block Diagram

3) LM35: To sense the external temperature we used LM35 module. It has temperature range of -55 to 150 degrees centigrade with an accuracy of 0.5 degree centigrade. It operates from 3V to 30V voltage range.

4) DS18B20: Water Temperature indicates how water is hot or cold. To measure temperature we are using waterproof DS18B20 temperature sensor which has are range from -55 to +125 degrees. This temperature sensor is digital type which gives accurate reading.

5) Microsoft Azure: Azure is a cloud platform created by Microsoft for deploying and testing applications. Azure has an IoT portal which helps to transfer sensor values from gateway to websites or database in the form of JSON packets.

6) PowerBI: Power BI is a business analytics service provided by Microsoft. It provides interactive visualizations where end users can create reports and dashboards. We used powerBI for visualization and analytic part of our project.

VII. Results and Analysis

In this project, the implemented several sensors that are used to determine the contamination of water and the results were produced. We used low cost sensors and hence the implementation cost for this system is minimal. Moreover, efficient algorithms are also being used here to co-ordinate the information collected by the sensors and hence there is no data loss and also the efficiency of transmission is increased. Whenever the quality of water is below threshold an email will be sent alerting concerned authorities.

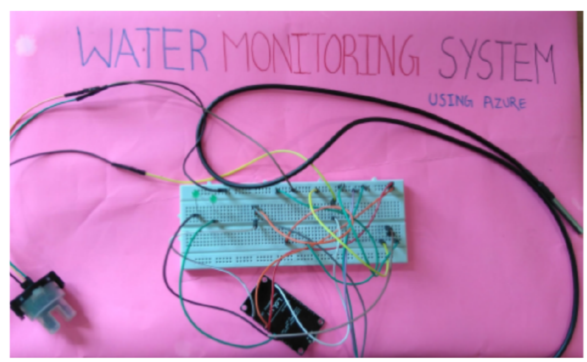

Fig. 3. Hardware Components

Fig 3 shows the complete hardware setup of the project, which includes NodeMCU, turbidity sensor, DS18B20 and LM35 temperature sensors connected through breadboard.

Fig 4 shows the log in screen of the web portal. It consists of a text box for user name and for password. On entering correct combination, authorized users can enter into the system.

Fig 5 shows the web page visualization of sensor data. Microsoft powerBI service is used to display the data in the

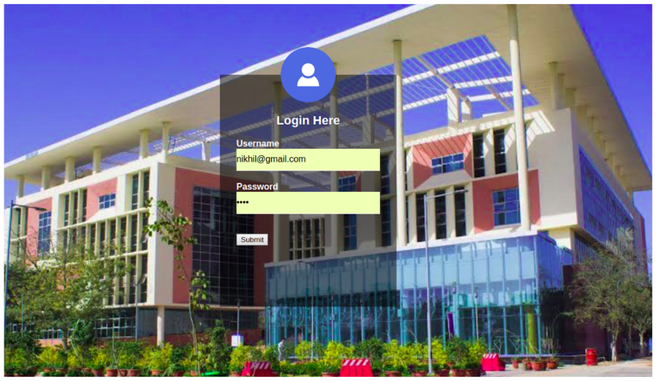

Fig. 4. Login portal of Website

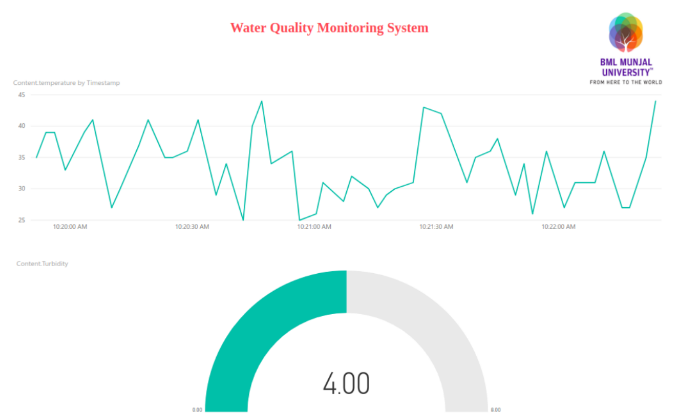

Fig. 5. Visualization of Sensor Data

form of graphs and charts. The displayed data will update automatically 5 times every day providing an updated information to users.

VIII. Comparison with existing research

All the previous work on water monitoring system has focused solely on measuring the quality of water using sensors such as pH and turbidity. Some of them have included cloud based module into their system for online view of sensor data. This paper also focuses on usability besides water quality monitoring. For usability we are controlling water temperature so as the users get water suitable to the climatic conditions. This is done by predicting external temperature using machine learning algorithms. All the previous models in this field have used heavy computing devices such as Arduino UNO and Raspberry Pie, whereas this paper uses NodeMCU which is light weight and also cost effective. This paper also includes Email alerts and Real time access to data visualization for authorized users.

IX. Conclusion

This paper presented a practical and economical solution to monitor the quality of water especially in rural areas without any human intervention. To solve this problem this paper presented various contemporary technologies such as IoT, cloud computing and Machine learning. On combining these technologies we are able to solve one of the basic and emerging problem of human survival to certain extent.

X. Future Work

- On integrating this system with state and central government work flow, we can enable fast response rate from government officers thus improving the quality of life in rural as well as urban areas.
- Since most of the villages in India doesn't have WiFi access we can add a Mobile GPRS module to our system enabling it to transmit data over 3G or 4G channels.
- Adding more quality sensors which can detect other chemical and physical parameters effecting the quality of water can improve our results and thus making our system effective.

References

[1] Pradeepkumar M, Monisha J. "The Real Time Monitoring of Water Quality in IoT Environment" 2016 International Journal of Innovative Research in Science, Engineering and Technology, 2015 ISSN(Online) : 2319-8753

[2] Atif Alamri, Wasai Shadab Ansari, Mohammad Mehedi Hassan, M. Shamim Hossain, Abdulhameed Alelaiwi, and M.Anwar Hossain, "A Survey on Sensor-Cloud: Architecture, Applications, and Approaches", International Journal of Distributed Sensor Networks, Volume 2013

[3] Kedia, Nikhil. "Water Quality Monitoring for Rural Areas- a Sensor Cloud Based Economical Project" 2015 1st International Conference on Next Generation Computing Technologies (NGCT), 2015, doi:10.1109/ngct.2015.7375081.

[4] Vijayakumar, N, and R Ramya. "The Real Time Monitoring of Water Quality in IoT Environment" 2015 International Conference on Circuits, Power and Computing Technologies [ICCPCT-2015], 2015, doi:10.1109/iccpct.2015.7159459.

[5] R.Karthik Kumar, M.Chandra Mohan, S.Vengateshapandiyan, M.Mathan Kumar, R.Eswaran. "Solar based advanced water quality monitoring system using wireless sensor network" 2014, International Journal of Science, Engineering and Technology Research, 2014

[6] Fiona Regan, Antoin, McCarthy. "Smart Coast Projectâ Smart Water Quality Monitoring Systemâ 2006, Marine Institute/Environmental Protection Agency Partnership, 2006

[7] Vaishnavi V. Daigavane, Dr. M.A Gaikwad. "Water Quality Monitoring System Based on IOT" 2017 Advances in Wireless and Mobile Communications, Nov 2017 ISSN 0973-6972

[8] "Resources." Resources — Internet of Things World Forum, www.iotwf.com/resources/72

[9] "Average Monthly Temperatures across the World." DataMarket, datamarket.com/data/set/1loo/average-monthly-temperatures-across-the-world!ds=1loo!1n6s=1nq.3jsdisplay=line.