IoT based Smart Water Quality Monitoring System

Monira Mukta
Department of Computer Science and Engineering
East West University
Dhaka, Bangladesh
e-mail: monira.mukta7@gmail.com

Samia Islam
Department of Computer Science and Engineering
East West University
Dhaka, Bangladesh
e-mail: islamsamia94@gmail.com

Surajit Das Barman
Department of Computer Science and Engineering
East West University
Dhaka, Bangladesh
e-mail: surajitbarman012@ewubd.edu

Ahmed Wasif Reza
Department of Computer Science and Engineering
East West University
Dhaka, Bangladesh
e-mail: wasif@ewubd.edu

M Saddam Hossain Khan
Department of Computer Science and Engineering
East West University
Dhaka, Bangladesh
e-mail: saddam.cse@ewubd.edu

Abstract—**This paper represents an IoT (Internet of things) based smart water quality monitoring (SWQM) system that aids in continuous measurement of water condition based on four physical parameters i.e., temperature, pH, electric conductivity and turbidity properties. Four sensors are connected with arduino-uno in discrete way to detect the water parameters. Extracted data from the sensors are transmitted to a desktop application developed in .NET platform and compared with the WHO (World Health Organization) standard values. Based on the measured result, the proposed SWQM system can successfully analyze the water parameters using fast forest binary classifier to classify whether the test water sample is drinkable or not.**

Keywords-internet of things; pH; electric conductivity; turbidity; arduino-uno; fast forest binary classifier

I. INTRODUCTION

The impact of water on any living beings is beyond description. With the rapid increase of world population, water management becomes an important issue specially in industrial, agricultural and other sectors. Most of the people around the world lack behind drinkable water. Every year many people are suffering from various fatal diseases caused by water pollution. Research has found that around 5 million death is caused only because of drinking unsafe water. Research by WHO (World Health Organization) shows that almost 1.4 million of child death can be prevented by providing drinkable water to them [1]. The primary objective of this project is to introduce an intelligent water quality monitoring system in IoT (Internet of Things) platform which would help to monitoring different physical parameters of the drinkable water rather than relying on manual process. Moreover, IoT is a system of alliance among various devices and the competence of deportation data over the system [2, 3].

Several research works have been conducted in recent times to develop intelligent system to identify and monitor water parameters. For real time monitoring of water quality and delivery, an in-pipe monitoring system based on sensor nodes is proposed [4]. Their proposed architecture focused on the low cost, lightweight implementation, pipeline electrochemical system and the sensors that are used for this architecture are optical sensors. This system is appropriate for large amount categorizations enabling an approach to water purchaser, water distributers and water supremacies. Authors in [5] has developed a broker-less architecture framework for both publisher and subscriber for monitoring water quality. They analyzed the measured data of temperature, pH and dissolved oxygen from water samples and results an inversely proportional relationship among them. An IoT based remote sensing system is introduced for collecting, monitoring and analyzing water quality in remote area for Fizi [6]. In article [7], a smart IoT based technology is explained for real time water quality monitoring system. An industrial water quality monitoring system using four different sensors e.g. turbidity, pH, temperature and level of water is developed in [8].

The goal of this research is to develop a smart water quality monitoring (SWQM) system using the IoT platform. Four physical parameters: temperature, pH, conductivity and turbidity of different water samples are measured via four sperate sensors equipped with Arduino Uno. The extracted sensor data are analyzed using the fast forest binary classifier.

978-1-7281-1322-7/19/$31.00 ©2019 IEEE

A desktop application is developed in .NET platform to identify whether the tested water samples are safe or unsafe for human consumption. The overview of the proposed SWQM system is introduced in Section II. Section III presents the design methodology including the hardware set-up and desktop application development. The measured data of different sensors and their analysis are described in Section IV. Lastly, Section V is concluded the paper.

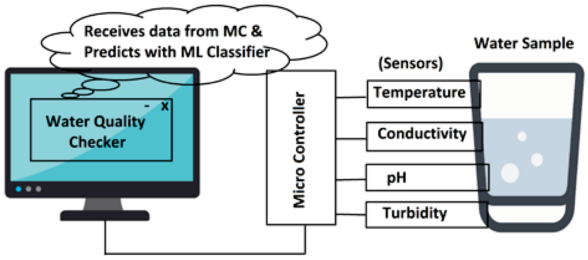


Figure 1. Block diagram of the proposed SWQM system.

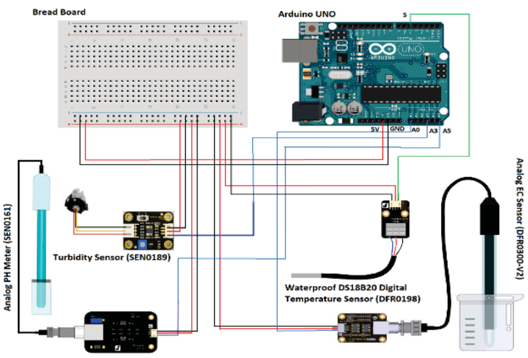


Figure 2. Circuit diagram of the hardware of SWQM system.

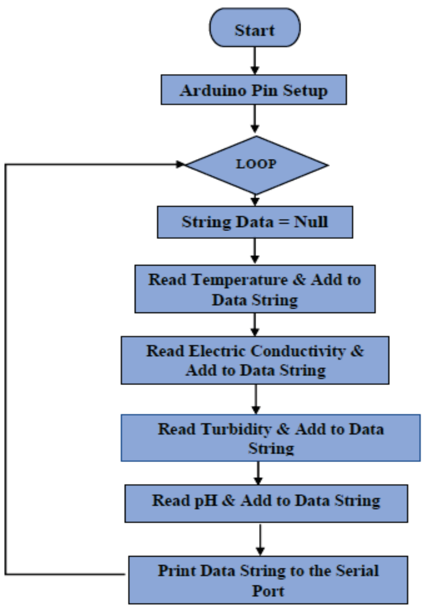


Figure 3. Flowchart of the working procedure of hardware.

II. System Overview

The proposed SWQM system is able to read data from water samples by sensors through the microcontroller and analyze them using machine learning algorithm to predict water quality. The proposed block diagram of SWQM system in Fig. 1 consists of four different sensors connected with controller to measure four important physical parameters (pH, temperature, conductivity and turbidity) of water samples. The pH sensor SEN0161 is used to measure the presence of acidity or alkalinity of any solution in logarithmic scale. The digital temperature sensor DFR0198 provides accurate reading between –55 to 125ºC. To measure the electrical conductivity of water sample, the analog sensor DFR0300 is utilized. The recommended detection range of this sensor is 1 to 15 ms/cm within a temperature between 0-40ºC. Turbidity sensor SEN0189 is used in the design to detect the presence of suspended particles by using light. The extracted data from these sensors are accessed by the controller arduino-uno and transfer them to the developed desktop application. Machine learning algorithm is implemented at the backend to predict the water quality based on the measured data. Since the system will predict either the test water sample is "Drinkable" or "Not Drinkable", the fast forest binary classifier algorithm is employed. 60 different water samples have been collected from nearby tap, filter, soft drinks and other sources. The prediction accuracy of the designed system is compared for the experimented data.

III. Design and Experiment

A. Circuit Diagram

Fig. 2 shows the schematic circuit diagram of the hardware set-up of the proposed SWQM system. Except the temperature sensor, other three sensors are of analog type. Each sensor has three different color wires such as red, black and others. Here, red wires are for +5V power supply, black wires are for ground and others are used for data estimation. A breadboard is used for creating common points for ground and power supply separately. Then common node of ground is connected to the ground of arduino and same process is repeated for power supply. The analog sensors are connected to the analog pins and digital sensor is connected to digital pin of the controller. The flow chart of the working procedure of controller and sensors is depicted in Fig. 3.

B. Machine Learning Algorithm

For the proposed SWQM system, the extracted sensor data are analyzed accordingly to predict the system's accuracy. Fast forest binary classifier has deployed here where different water samples i.e. salt, mud, drain, tap, soft drinks and drinking water are taken for training the data set. The average combination of many small and weak decision trees in fast forest regression model forms a strong learner. The algorithm works as follow: for each tree in the forest a bootstrap sample from *Z* are selected. At that situation *Z* is the i^{th} bootstrap. The method of this decision tree is modified by the decision tree learning algorithm. Randomly subset of features of $a \subseteq X$, where X denotes the set of features. Here, a is so smaller than X. The fast forest algorithms return R which is the set of train model. The narrowing of the set of features, the learning set of features is speed up drastically. The pseudocode of the proposed fast forest binary algorithm is given below as [9].

Algorithm Fast Forest

Precondition: A training set Z := $(x_{11}, x_{12}, x_{13}, y_{11}), \ldots, (x_{n1}, x_{n2}, x_{n3}, y_{n1})$, features X, and number of trees in forest T.

function FastForest(Z, X)
  $R \leftarrow \emptyset$
  loop $i \in 1, \ldots, T$ *do*
    $Z^{(i)}$ $\leftarrow$A bootstrap sample from Z
    r_i $\leftarrow$ LearnTreeModel ($Z^{(i)}$, X)
    $R \leftarrow R \cup \{r_i\}$
  end loop
  return R
end function
function LearnTreeModel(Z, X)
  At each node:
    $a \leftarrow$ very small subset of X
    Split on best feature in a
  return Learned tree model
end function

C. Developed Desktop Application

The performance of the adopted fast forest binary classifier is compared with three other binary classifiers: support vector machine (SVM), logistic regression and average perceptron techniques. Among four algorithms, fast forest binary classifier provides better accuracy for the same set of data and used to develop the desktop application "Sprinkle: Water Quality Checker" (Fig. 4) for monitoring the water quality. Fig. 5, demonstrates the working scheme of the desktop application built in .NET platform. Firstly, ports connected with the arduino are selected. Then, data are read with the assistance of the sensors. These data are used to check whether the water sample is drinkable or not drinkable, and the result is saved into the database. During the processing of data, only three parameters (pH, Conductivity and Turbidity) are considered, because temperature is used in the experiment as a factor of conductivity. The complete experimental set-up of the developed SWQM system is shown in Fig. 6.

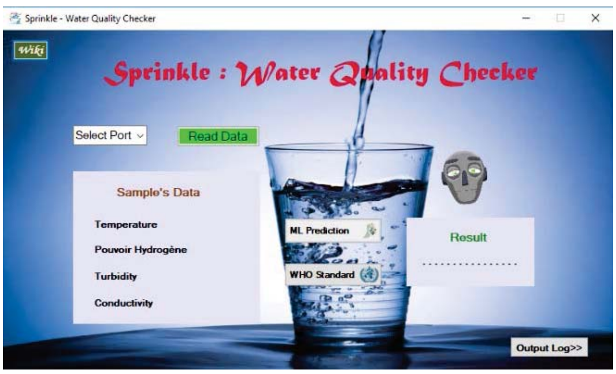


Figure 4. Sprinkle: Water quality checker for SWQM system.

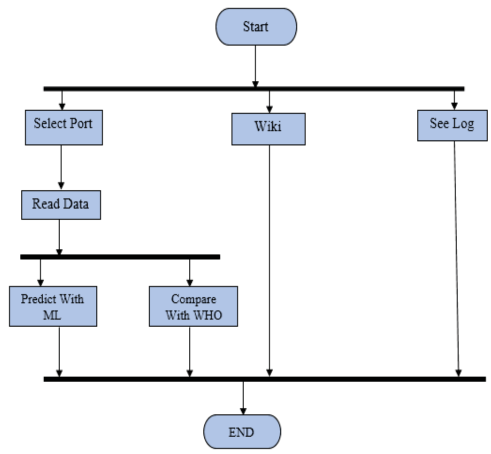
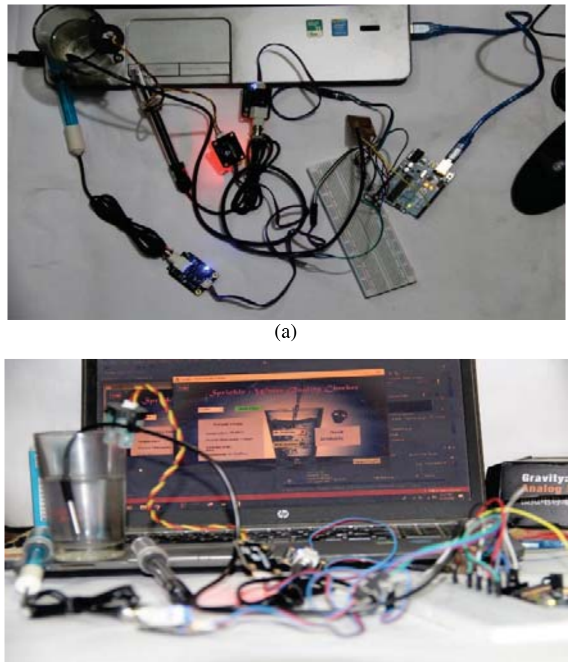


Figure 5. Flow chart of the developed desktop application for SWQM system.



(a)

(b)

Figure 6. Experimental setup of SQWM system.

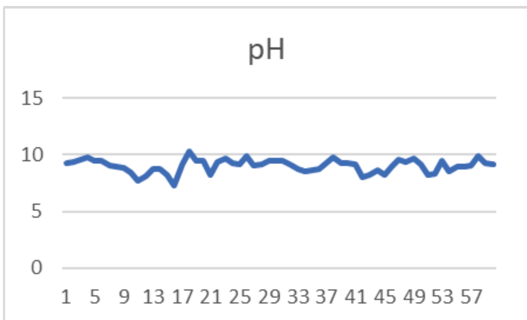


Figure 7. Collected data for PH parameter.

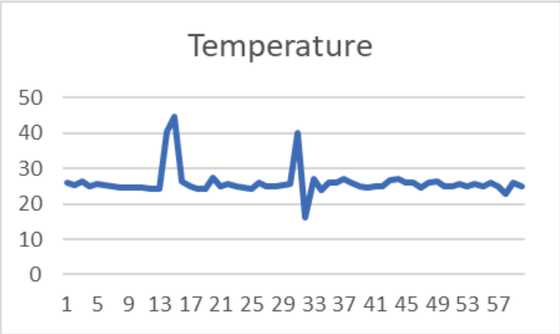


Figure 8. Collected data for temperature.

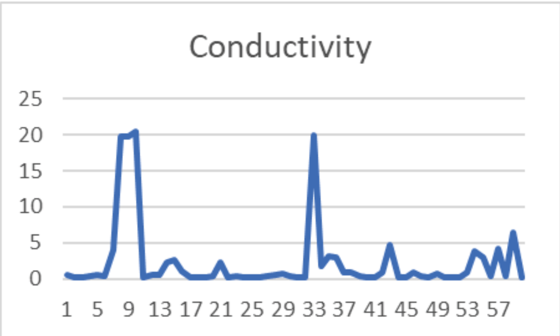


Figure 9. Collected data for conductivity.

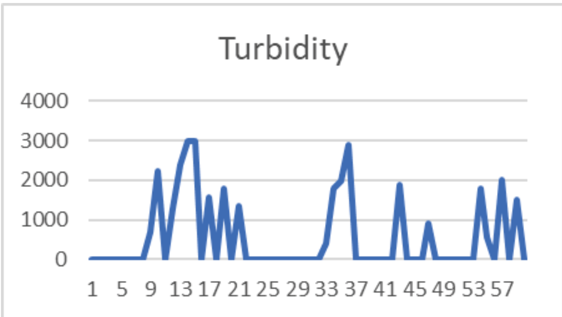


Figure 10. Collected data for turbidity.

IV. Result and Discussion

Sixty water samples are collected from different water sources, and tested to measure the parameters i.e pH, temparature, electric conductivity and turbidity for each sample. These water sources are divided into three catagories: natural, impure and potable water sources. Fig. 7 to 10 illustrates the extracted values corresponding four physical parameters for each water sample. Table I shows the analysis of the each physical parameter for 60 samples (83% of data were used for training and 17% were used as test data) divided into three catagories of water sources and indicates their percentage of concentrations not within the WHO standard values [10]. Table II shows the experimental values of detected parameters and system's prediction for different water samples.

TABLE I. Analysis of Physical Calculation

Parameters	Water Source	N	Min.	Med.	Max.	WHO SV	% not within SV
Temperature	Natural	23	24.11	25.11	27.2		-
	Impure	16	16	24.6	44.63		-
	Potable	21	23	25.76	27		-
	All	60	16	25.09	44.63		-
Conductivity	Natural	23	0.12	0.34	6.45	0.3-0.8 mS/cm	69.57
	Impure	16	0.12	0.57	20.44		81.25
	Potable	21	0.12	0.48	19.89		66.67
	All	60	0.12	0.44	20.44		71.67
PH	Natural	23	7.3	9.23	9.88	6.5-8.5	82.61
	Impure	16	7.66	8.9	10.26		68.75
	Potable	21	7.99	9.12	9.89		85.71
	All	60	7.3	9.15	10.26		80
Turbidity	Natural	23	0	0	2023	<5 NTU	13.04
	Impure	16	0	989.58	3000		56.25
	Potable	21	0	0	2902.89		33.33
	All	60	0	0	3000		31.67

TABLE II. Experimental Values of Detected Parameters and System's Prediction

Water Sample	Temparature	EC	pH	Turbidity	System's prediction
1	25	0.14	8.2	0	Drinkable
2	25.67	0.19	8.3	0	Drinkable
3	25.03	0.89	9.5	0	Drinkable
4	25.76	3.89	8.5	1791	Not Drinkable
5	24.99	2.9	8.9	566	Not Drinkable
6	26	0.31	9	0	Drinkable
7	25	4.24	9.1	2023	Not Drinkable
8	23	4.39	9.89	0	Drinkable
9	25.99	6.45	9.29	1520	Not Drinkable
10	24.89	0.18	9.11	0	Drinkable

The general guideline for pH level in drinking water is around 6.5-8.5 suggested by WHO. From the table it is observed that, almost 80% of the tested water samples are beyond the recommended pH range, being alkaline in nature which indicates the presence of carbonates and limestone in the water samples. In consequence, excess presence of alkanity in human body can cause skin irritation, gastrointestinal and metabolic alkalosis. Turbidity represents a key issue in terms of analysing the microorganism quality of water. According to the guidline, the acceptable turbidity level should be below 5 NTU. Results shows the high value of turbidity for impure water compare to the natural water samples. In case of measure the water quality or pollution level, electric conductivity plays a vital role. Drinkable water conductivity ranging from 0.3-0.8 μS/cm. According to the table, more than 70% of the test sample's conductivity is beyond the WHO standard value.

TABLE III. CONFUSION MATRIX

Actual	Total = TP+TN+FP+FN	Prediction	
		Positive	Negative
	Positive	TP	FN
	Negative	FP	TN

To verify the system's performance the measured data from the sensors are analyzed by using fast forest binary classifier, and compared with mentioned three other algorithms in terms of F1score and accuracy shown in Fig. 11. Table III shows the confusion matrix to represent the relation between the actual and predicted class. From the table, the accuracy of the system can be predicted by using (1).

$$\text{Accuracy} = (TP + TN)/\text{Total} \quad (1)$$

To determine the predictive power of the classifiers, the F1score is calculated via (2).

$$\text{F1score} = 2 \times [(PR \times RE)/(PR + RE)] \quad (2)$$

Here, PR= precision, RE=recall. Table IV shows the F1score and accuracy of four algorithms where fast forest binary classifier provides highest accuracy rate is 100% and has been adopted to develop the application Sprinkle: Water Quality Checker (Fig. 12) to predict the test water samples either "Drinkable" or "Not Drinkable".

TABLE IV. ACCURACY COMPARISON

Binary Classifiers	F1score	Accuracy
Logistic Regression	75.00%	80.00%
SVM	75.00%	80.00%
Averaged Perceptron	75.00%	80.00%
Fast Forest	100%	100%

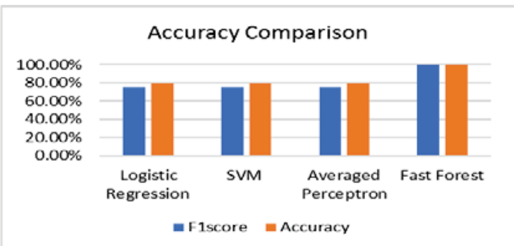


Figure 11. Accuracy comparison for four algorithms.

V. CONCLUSION

The ultimate goal of this work is to observe the quality of water samples by designing a smart water quality monitoring (SWQM) device implemented in IoT platform that can detect four specific physical parameters: temperatures, pH, turbidity and conductivity in water, and analyze the extracted value of these parameters using suitable machine learning approach. Different water samples are tested with the assistance of arduino based sensors and collected their values of different metrics. Fast forest binary classifier shows better scrutinizing performance to validate the system's accuracy and effectiveness in predicting water quality. The SWQM system has shown its importance by providing accurate performance detecting the water quality based on physical parameters. With the upgrade feature of IoT technology to detect chemical parameters of water, this system can be implemented for real time water monitoring solution in near future.

Figure 12. Prediction of SWQM system for water sample.

REFERENCES

[1] World Health Organization, WHO/UNICEF Joint Water Supply, and Sanitation Monitoring Programme. Progress on sanitation and drinking water: 2015 update and MDG assessment. World Health Organization, 2015.

[2] R. Ramakala, S. Thayammal, A. Ramprakash and V. Muneeswaran. "Impact of ICT and IOT Strategies for Water Sustainability: A Case study in Rajapalayam-India," Int. Conf. Computational Intelligence and Computing Research (ICCIC 18), IEEE, Dec. 2017, pp. 1-4, doi:10.1109/ICCIC.2017.8524399.

[3] N. Zidan, M. Mohammed and S. Subhi, "An IoT based monitoring and controlling system for water chlorination treatment," Proc. Int. Conf. Future Networks and Distributed Systems, ACM, Jun. 2018, p. 31, doi:10.1145/3231053.3231084.

[4] J. Ramprabu and C. Paramesh, "Automated sensor network for monitoring and detection of impurity in drinking water system," International Journal for Research in Applied Science & Engineering Technology, vol. 3, Jan. 2015, pp. 275-280.

[5] A. A. Pranata, M. L. Jae and S. K. Dong, "Towards an IoT-based water quality monitoring system with brokerless pub/sub architecture," Int. Symp. Local and Metropolitan Area Networks (LANMAN 17), IEEE, Jun. 2017, pp. 1-6, doi:10.1109/LANMAN.2017.7972166.

[6] A. N. Prasad, K. A. Mamun, F. R. Islam and H. Haqva, "Smart water quality monitoring system," Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE), IEEE, Dec. 2015, pp. 1-6,

[7] S. Geetha and S. Gouthami, "Internet of things enabled real time water quality monitoring system," Smart Water, vol. 2, 2016, pp. 1, doi:10.1186/s40713-017-0005-y.

[8] S. M. G. Nikkam and V. R. Pawar, "Analyzing water quality for industrial application under IOT environment," International Research Journal of Engineering and Technology, vol. 3, 2016, pp. 882-885.

[9] Random Forest – Wisc. Available- http://pages.cs.wisc.edu/~matthewb/pages/notes/pdf/ensembles/RandomForests.pdf

[10] WHO, Guidelines for Drinking-Water Quality. 4th ed.; World Health Organisation: Geneva, Switzerland, 2012.