Real Time – River Water Quality Monitoring and Control System

TEAM ID – PNT2022TMID03218

Team Leader: Velmurugan R

Team Member 1: Vallimaharajan G

Team Member 2: Varthayini R

Team Member 3: Vengadesh K

Team Member 4: Sruthi R

Abstract

Current water quality monitoring system is a manual system with a monotonous process and is very time-consuming. This paper proposes a sensor-based water quality monitoring system. The main components of Wireless Sensor Network (WSN) include a microcontroller for processing the system, communication system for inter and intra node communication and several sensors. Real-time data access can be done by using remote monitoring and Internet of Things (IoT) technology. Data collected at the apart site can be displayed in a visual format on a server PC with the help of Spark streaming analysis through Spark MLlib, Deep learning neural network models, Belief Rule Based (BRB) system and is also compared with standard values. If the acquired value is above the threshold value automated warning SMS alert will be sent to the agent. The uniqueness of our proposed paper is to obtain the water monitoring system with high frequency, high mobility, and low powered. Therefore, our proposed system will immensely he lp Bangladeshi populations to become conscious against contaminated water as well as to stop polluting the water.

1.Introduction

The environment around consists of five key elements e.g., soil, water, climate, natural vegetation, and landforms. Among these water is the utmost crucial element for human life. It is also vital for the persistence of other living habitats [1]. Whether it is used for drinking, domestic use, and food production or recreational purposes, safe and readily available water is the need for public health [2]. So it is highly imperative for us to maintain water quality balance. Otherwise, it would severely damage the health of the humans and at the same time affect the ecological balance among other species [3]. Water pollution is a foremost global problem which needs ongoing evaluation and adaptation of water resource directorial principle at the levels of international down to individual wells. It has been studied that water pollution is the leading cause of mortalities and diseases worldwide. The records show that more than 14,000 people die daily worldwide due to water pollution. In many developing countries, dirty or contaminated water is being used for drinking without any proper prior treatment. One of the reasons for this happening is the ignorance of public and administration and the lack of water quality monitoring system which makes serious health issues [3, 4]. In this paper, we depict the design of Wireless Sensor Network (WSN) [4-7] that assists to monitor the quality of water with the support of information sensed by the sensors dipped in water. Using different sensors, this system can collect various parameters from water, such as pH, dissolved oxygen, turbidity, conductivity, temperature, and so on. The rapid development of WSN technology provides a novel approach to real-time data acquisition, transmission, and processing. The clients can get ongoing water quality information from far away. Now a day's Internet of things (IoT) is an innovative technological phenomenon. It is shaping today's world and is used in different fields for collecting, monitoring and analysis of data from remote locations. IoT integrated network if everywhere starting from smart cities, smart power grids, and smart supply chain to smart wearable [7- 12]. Though IoT is still under applied in the field of environment it has huge potential. It can be applied to detect forest fire and early earthquake, reduce air population, monitor snow level, prevent landslide, and avalanche etc. Moreover, it can be implemented in the field of water quality monitoring and controlling system [4, 13]. Water quality monitoring has gained more interest among researchers in this twenty-first century. Numerous works are either done or ongoing in this topic focusing on various aspects of it. The key theme of all the projects was to develop an efficient, cost-effective, real-time water quality monitoring system which will integrate wireless sensor network and internet of things [14]. In this research, we monitor the physical and chemical parameters of water bodies inside Chittagong city by using an IoT based sensor network.

2. Related Work

Figure 1 shows the general building blocks of smart online monitoring solutions considered in this section. Three main subsystems identified include

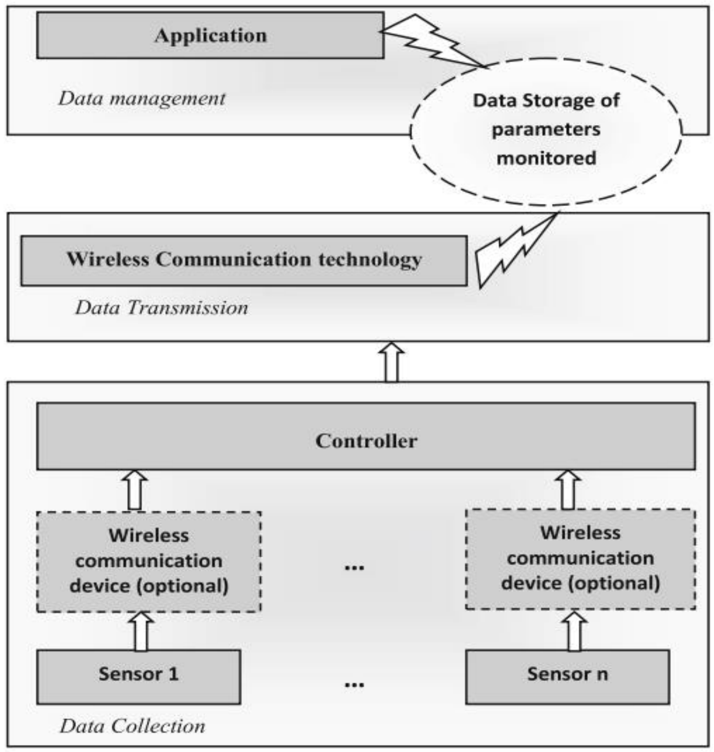


Fig. 1 Block diagram of Smart water Quality monitoring

Data management subsystem includes the application which accesses the data storage cloud and displays the same to the end user. Data transmission subsystem consists of a wireless communication device along with build in security features, which transmits the data from the controller to data storage cloud. Data collection subsystem consists of multi-parameter sensors and optional wireless communication device to transmit the sensor information to the controller. A controller gathers the data, processes the same. Sensors form the bottom most part of the block diagram. Several sensors are available to monitor water quality parameters. These sensors are placed in the water to be tested which can be either stored water or running water. Sensors convert the physical parameter into equivalent measurable electrical quantity, which is given as input to controllers through an optional wireless communication device. Main function of the controller is to read the data from the sensor, optionally process it, and send the same to the application by using appropriate communication technology. Choice of the communication technology and the parameters to be monitored depends on the need of the application. Application includes the data management functions, data analysis and alert system based on the monitored parameters. This section further discusses the previous work carried out in each of the subsystems. Application Online smart water quality has been proposed for several applications in literature as shown in Table 1. Domestic water is intended for human consumption for drinking and cooking purposes. The Bureau of Indian Standards (Central Ground Water Board, 2017) provides details about acceptable limits of substances such as Aluminium, Ammonia, Iron, Zinc etc. Traditional water quality measurement involves manual collection of water at

Application	References
Domestic running water	Vijayakumar and Ramya (2015), Niel et al. (2016), Theofanis et al. (2014), Jayti and Jignesh (2016), Poonam et al., 2016, Xin et al. (2011), Xiuli et al. (2011), Offiong et al. (2014)
Domestic Stored water	Thinagaran et al. (2015), Vinod and Sushama (2016), Pandian and Mala (2015), Azedine et al. (2000), Sathish et al. (2016)
Lake, River, Sea water, Environmental monitoring	Tomoaki et al. (2016), Vinod and Sushama (2016), Peng et al. (2009), Francesco et al. (2015), Christie et al. (2014), Haroon and Anthony (2016), Anthony et al. (2014), Li et al. (2013)
Aquaculture centers	Goib et al. (2015), Xiuna et al. (2010), Gerson et al. (2012)
Drinking water distribution systems	Eliades et al. (2014), Ruan and Tang (2011)
Water and Air quality	Mitar et al. (2016)
Not limited to specific application	Liang (2014), Wei et al. (2012)

various locations, storing the samples in centralized location and subjecting the samples to laboratory analytical testing (Thinagaran et al., 2015; Vinod & Sushama, 2016; Pandian & Mala, 2015; Azedine et al., 2000; Offiong et al., 2014). Such approaches are not considered efficient due to the unavailability of real time water quality information, delayed detection of contaminants and not cost effective solution. Hence, the need for continuous online water quality monitoring in highlighted in (Vijayakumar & Ramya, 2015; Niel et al., 2016; Theofanis et al., 2014; Bhatt & Patoliya, 2016; Poonam et al., 2016; Xin et al., 2011; Xiuli et al., 2011; Sathish et al., 2016). Smart water quality approaches have been considered for lake and sea water applications. For such applications, distributed wireless sensor networks are required to monitor the parameters over a larger area and send the data monitored to a centralized controller using wireless communication. Such applications normally monitor parameters such as chlorophyll (Francesco et al., 2015), dissolved oxygen concentration (Christie et al., 2014; Anthony et al., 2014) and temperature (Peng et al., 2009; Francesco et al., 2015; Christie et al., 2014). Aquaculture centers require water quality monitoring and forecasting for healthy growth of aquatic creatures (Goib et al., 2015; Gerson et al., 2012; Xiuna et al., 2010). In (Gerson et al., 2012) authors have developed biosensors using Arduino microcontroller to monitor animal behavioral changes due to aquatic pollution. The abnormal behavior of animals can be considered as an indication of water contamination. In (Xiuna et al., 2010) authors have proposed a smart water quality monitoring system to forecast water quality using artificial neural networks. Extensive tests have been carried out for a period of 22 months at isolated local area network and the data has been transferred to internet using CDMA technology. Water quality monitoring in distribution systems is challenging in the context of management of distributed wireless sensor networks (WSN). A water distribution network for monitoring chlorine concentration has been presented in (Eliades et al., 2014). Solar enabled distributed WSN has been proposed in (Ruan & Tang, 2011) for monitoring parameters such as pH, turbidity and oxygen density. Water at different sites is monitored in real-time using an architecture composed of solar cell enabled sensor nodes and base station. Flexibility, low carbon emission and low power consumption are the advantages of the method proposed in the paper. A combined system for water and air quality measurement is proposed in (Mitar et al., 2016) using additional sensors for measuring air temperature and relative humidity. Parameters monitored Based on extensive experimental evaluation carried out by US Environmental Protection Agency (USEPA) it has been concluded that chemical and biological contaminants used have an effect on many water parametersrs monitored including Turbidity (TU), Oxidation Reduction Potential (ORP), Electrical Conductivity (EC) and pH. Thus, by monitoring and detecting changes in the water parameters, it is feasible to infer the water quality (Theofanis et al., 2014). A detailed list of work carried out to monitor water parameters is presented in Table 2. The pH of the water is one of the most important factors when investigating water quality, as it measures how basic or acidic the water is. Water with a pH of 11 or higher can cause irritation to the eyes, skin and mucous membrane. Acidic water (pH 4 and below) can also cause irritation due to its corrosive effect (Niel et al., 2016). Measurement of dissolved oxygen (DO) is important for aquaculture centers since this parameter determines whether or not a species can survive in the said water source. ORP is a measure of degree to which a substance is capable of oxidizing or reducing another substance. ORP is measured in milli volts (mv) using an ORP meter. Tap water and bottled water have a positive value of ORP. Turbidity refers to concentration of suspended particles in water. Conductivity gives an indication of the amount of impurities in the water, the cleaner the water, the less conductive it is. In many cases, conductivity is also directly associated with the total dissolved solids (TDS).

Parameters monitored	References
pH	Vijayakumar and Ramya (2015), Mitar et al. (2016), Tomoaki et al. (2016), Vinod and Sushama (2016), Niel et al. (2016), Goib et al. (2015), Theofanis et al. (2014), Peng et al. (2009), Jayti and Jignesh (2016), Poonam et al., 2016, Xin Wang (2011), Gerson et al. (2012), Pandian and Mala (2015), Liang (2014), Xiuna et al. (2010), Christie et al. (2014), Azedine et al. (2000), Offiong et al. (2014), Anthony et al. (2014), Sathish et al. (2016)
Dissolved Oxygen	Vijayakumar and Ramya (2015), Goib et al. (2015), Jayti and Jignesh (2016), Gerson et al. (2012), Liang (2014), Xiuna et al. (2010), Christie et al. (2014), Offiong et al. (2014), Anthony et al. (2014)
Oxidation reduction potential	Niel et al. (2016), Theofanis et al. (2014)
Temperature	Vijayakumar and Ramya (2015), Mitar et al. (2016), Niel et al. (2016), Theofanis et al. (2014), Peng et al. (2009), Jayti and Jignesh (2016), Poonam et al., 2016, Gerson et al. (2012), Pandian and Mala (2015), Liang (2014), Xiuna et al. (2010), Francesco et al. (2015), Christie et al. (2014), Azedine et al. (2000), Anthony et al. (2014)
Turbidity	Vijayakumar and Ramya (2015), Tomoaki et al. (2016), Vinod and Sushama (2016), Theofanis et al. (2014), Jayti and Jignesh (2016), Poonam et al., 2016, Gerson et al. (2012), Pandian and Mala (2015), Francesco et al. (2015), Offiong et al. (2014), Sathish et al. (2016)
Conductivity	Vijayakumar and Ramya (2015), Niel et al. (2016), Theofanis et al. (2014), Jayti and Jignesh (2016), Gerson et al. (2012), Francesco et al. (2015), Christie et al. (2014), Azedine et al. (2000), Anthony et al. (2014), Sathish et al. (2016)
Water level sensing	Thinagaran et al. (2015)
Flow sensing	Niel et al. (2016)
Air temperature	Mitar et al. (2016)
Relative Humidity	Mitar et al. (2016)
Presence of organic compounds	Mitar et al. (2016)
Chlorine concentration	Eliades et al. (2014), Francesco et al. (2015)
Chlorophyll	Francesco et al. (2015)

Communication technology used Wireless technology is used for communication between sensor to controller and from controller to data storage cloud as shown in Fig. 1. Different technology has been used in each of the communication scenario. Table 3 shows the frequently used wireless communication technology for information transfer. Communication between sensors and controller Sensors are connected to the controller, either directly using UART protocol or remotely using Zigbee protocol. ZigBee is a technology of data transfer in wireless

Communication	Technology used	References
Between sensors and controller	Zigbee	Vinod and Sushama (2016), Niel et al. (2016), Theofanis et al. (2014), Peng et al. (2009), Jayti and Jignesh (2016), Poonam et al., 2016, Xin et al. (2011), Pandian and Mala (2015), Liang (2014), Christie et al. (2014), Offiong et al. (2014), Anthony et al. (2014), Li et al. (2013)
	UART	Tomoaki et al. (2016), Wei et al. (2012), Vijayakumar and Ramya (2015), Thinagaran et al. (2015), Mitar et al. (2016), Sathish et al. (2016)
Between controller and application	GSM/GPRS	Peng et al. (2009), Xin et al. (2011), Liang (2014), Wei et al. (2012), Francesco et al. (2015), Anthony et al. (2014), Tomoaki et al. (2016)
	Ethernet LAN	Theofanis et al. (2014)
	IoT (using external WiFi Module)	Vijayakumar and Ramya (2015), Thinagaran et al. (2015), Mitar et al. (2016), Jayti and Jignesh (2016), Poonam et al., 2016, Sathish et al. (2016)
	IoT (using inbuilt WiFi Module)	Proposed
	LCD, Alarm, Actuators.	Vinod and Sushama (2016), Niel et al. (2016), Li et al. (2013)

network. It has a low energy consumption and is designed for multichannel control systems, alarm system and lighting control. ZigBee builds on the physical layer and media access control defined in IEEE standard 802.15.4 for low-rate WPANs. In smart water quality systems, Zigbee protocol is used for communication between sensor nodes and the controller when the sensors are placed in remote location away from the controller. For in-pipe domestic monitoring, direct connection of sensors and controller is preferred. In (Tomoaki et al., 2016) authors have developed a WSN system for water quality monitoring. Sensors are connected to the transmission module using UART. Communication with the outside of the sensor nodes is performed with the Internet connection using the 3G mobile network. Authors in (Theofanis et al., 2014) have proposed a water quality monitoring system for in-pipe monitoring and assessment of water quality on fly. Sensor nodes are installed in the pipes that supply water at consumer sites. Communication between controller and data storage Communication between controller and centralized data storage is carried out using long range communication standards such as 3G and Internet. Some the previous works aim at alerting the user in form of SMS about the water quality. Such systems (Peng et al., 2009; Xin et al., 2011; Liang, 2014; Wei et al., 2012) require additional SIM card for the GPRS module connected with the controller. The drawbacks of such systems are additional cost for SIM card operation. Also, large quantities of data storage and retrieval are not possible at the user premises. Recently, IoT enabled solutions are gaining importance. Authors in (Alessio et al., 2016) provide a survey on the wide range of applications possible with Internet of Things and Cloud computing. IoT is a recent communication paradigm in which objects of everyday day life are equipped with microcontrollers, transceivers for digital communication, which will make the objects communicate with one another and the users, thus becoming an integral part of the Internet (Bushra & Mubashir, 2016; Biljana et al., 2017; Andrea et al., 2014). In (Vijayakumar & Ramya, 2015; Thinagaran et al., 2015; Mitar et al., 2016) an external Wi-Fi Table 3 Wireless communication technology used Communication Technology used References Between sensors and controller Zigbee Vinod and Sushama (2016), Niel et al. (2016), Theofanis et al. (2014), Peng et al. (2009), Jayti and Jignesh (2016), Poonam et al., 2016, Xin et al. (2011), Pandian and Mala (2015), Liang (2014), Christie et al. (2014), Offiong et al. (2014), Anthony et al. (2014), Li et al. (2013) UART Tomoaki et al. (2016), Wei et al. (2012), Vijayakumar and Ramya (2015), Thinagaran et al. (2015), Mitar et al. (2016), Sathish et al. (2016) Between controller and application GSM/GPRS Peng et al. (2009), Xin et al. (2011), Liang (2014), Wei et al. (2012), Francesco et al. (2015), Anthony et al. (2014), Tomoaki et al. (2016) Ethernet LAN Theofanis et al. (2014) IoT (using external WiFi Module) Vijayakumar and Ramya (2015), Thinagaran et al. (2015), Mitar et al. (2016), Jayti and Jignesh (2016), Poonam et al., 2016, Sathish et al. (2016) IoT (using inbuilt WiFi Module) Proposed LCD, Alarm, Actuators. Vinod and Sushama (2016), Niel et al. (2016), Li et al. (2013) Geetha and Gouthami Smart Water (2017) 2:1 Page 6 of 19 module is connected to the controller, which enables the controller to get connected to the nearest Wi-Fi hotspot and subsequently to the Internet cloud. Controller used Different controllers have been used in literature for smart water quality monitoring as listed in Table 4. Though each controller have its own salient features, most of the controllers used in literature work with external GPRS / Wi-Fi module for connectivity to the data storage or application. The proposed model in this paper uses TI CC3200, a controller with built in Wi-Fi module and dedicated ARM MCU for wireless communication purpose. TI CC3200 reduces the complexity and improves speed of operation compared with controllers with external Wi-Fi module (Texas instrument CC3200 Simple Link, 2017). A comparison of various microcontrollers and embedded boards used in the literature for smart water quality monitoring are provided in Table 7. Sensors used Several sensors are commercially available for water quality monitoring. Such sensors are used in (Thinagaran et al., 2015; Vinod & Sushama, 2016; Niel et al., 2016). Some of the works published in literature include fabricated sensors for improved usability. A fabricated buoy type sensor node is used for parameter monitoring in (Tomoaki et al., 2016). The fabricated sensor includes a solar cell, Li-ion battery, a power module and transmission module. A in-house fabricated TiO2 based thick film pH resistive sensor is used in (Mitar et al., 2016). This sensor module output can be directly connected to the microcontroller without additional signal processing electronics. In (Theofanis et al., 2014), authors have developed low cost, easy to use and accurate turbidity sensor for continuous in-pipe turbidity monitoring. In (Francesco et al., 2015), authors have presented a sea water probe for monitoring multiple parameters intended for sea water quality monitoring Table 5. Extensive data analysis and information processing has been presented in (Theofanis et al., 2014; Peng et al., 2009; Xiuna et al., 2010; Francesco et al., 2015; Azedine et al., 2000). A hierarchical routing algorithm to reduce the communication overhead and increase the life time of WSN suitable for river/lake water monitoring has been presented in (Haroon & Anthony, 2016). In (Public Utilities Board Singapore (PUB), 2016), a review on Smart Water Grid system with integration of communication technologies (ICT) is provided. An integrated

Controller used	References
AtMega	Thinagaran et al. (2015), Mitar et al. (2016)
PIC	Theofanis et al. (2014), Niel et al. (2016), Vinod and Sushama (2016)
Raspberry pi + IOT	Vijayakumar and Ramya (2015), Jayti and Jignesh (2016), Sathish et al. (2016)
ARM LPC	Francesco et al. (2015), Poonam et al. (2016)
Arduino	Anthony et al. (2014), Christie et al. (2014), Pandian and Mala (2015), Gerson et al. (2012)
8051	Li et al. (2013)
MSP430	Peng et al. (2009)
TI CC3200	Proposed work

Sensors used	References	Remarks
Fabricated buoy type sensor node.	Tomoaki et al. (2016)	Solar enabled sensor node with power module and transmission module
Solar cell enabled sensors	Ruan and Tang (2011)	
Fabricated TiO2-based thick film pH resistive sensor	Mitar et al. (2016)	Designed to ensure reliable measurements without any additional signal processing
Fabricated Turbidity sensor	Theofanis et al. (2014)	Designed to be compatible with WSN technology, in-pipe placement, low cost and accuracy
ISO/IEC/IEEE 21451–2 compliant sea water probe	Francesco et al. (2015)	Single Probe capable of measuring water temperature, salinity/ conductivity, turbidity and chlorophyll
Standard commercially available sensors	Vijayakumar and Ramya (2015), Thinagaran et al. (2015), Vinod and Sushama (2016), Niel et al. (2016), Theofanis et al. (2014), Eliades et al. (2014), Peng et al. (2009), Jayti and Jignesh (2016), Poonam et al. (2016), Xin et al. (2011), Gerson et al. (2012), Pandian and Mala (2015), Liang (2014), Wei et al. (2012), Xiuna et al. (2010), Francesco et al. (2015), Christie et al. (2014), Offiong et al. (2014), Anthony et al. (2014), Sathish et al. (2016), Li et al. (2013)	

management model covering the entire water cycle from sources to tap for securing the stability, safety and efficiency of water has been discussed in (Woon et al., 2016). Power consumption related issues Power consumption is a major constraint for IoT applications, because the applications are most likely to operate on batteries. Communication of data is a major source of power consumption. For applications such as smart water quality monitoring, data communication occurs in two stages. One is the communication between sensors and the controller and other is the communication between controller and application. Table 6 shows several possible short distance communication protocols applicable (Al-Fuqaha et al., 2015; Ray, 2016). Possible protocols for communication between sensor nodes and controller are Zigbee, Blue tooth, BLE and LoRa. Wi-Fi is not suitable for communication between sensor nodes and the controller because the power dissipation is high (Shuker et al., 2016). As per our literature survey, all the works have used zigbee protocol for communication between sensor nodes and controller.

Table 6 Short distance communication protocols

Parameter	Short distance protocols			
	ZigBee	Bluetooth	LoRa	Wi-Fi
Standard	IEEE 802.15.4	IEEE 802.15.1	LoRaWAN R1.0	IEEE 802.11 a/c/b/d/g/n
Transmission range	10–20 m	8–10 m	<30 Km	20–100 m
Energy consumption	Low	Bluetooth: Medium BLE: Very Low	Very low	High
Data Rate	40–250 Kb/s	1–24 Mb/s	0.3–50 Kb/s	1 Mb/s–6.75 Gb/s
Cost	Low	Low	High	High
Usage	Communication between sensor and controller			Communication between controller and application

The proposed work is aimed at domestic water quality monitoring. The sensors are assumed to be connected in-pipe. The controller and the sensors form a single module installed in the user premises. Therefore, the sensors are directly connected to controller. For applications such as lake, river and sea water monitoring, sensors and the controller are separated by considerable distance. Under such conditions, short range communication protocols (such as Zigbee), listed in Table 6 are used. For communication between controller and the application, Wi-Fi is a compelling choice. With other short range protocols, the sensor nodes communicate with controller easily, but when trying to connect the system to the Internet some type of adapter that is able to communicate with both the sensors and the Internet is needed. This is additional hardware overhead. With Wi-Fi, the above problem does not arise, because there is an infrastructure that is already built and is in existence. The limitation of WiFi is that, the standard was designed for laptops and PCs, where power requirement is completely different from battery operated smart objects. Hence, manufacturers have started developing low-power Wi-Fi devices. Power management and extended battery life are primary focus areas for embedded low-power Wi-Fi devices such as CC3200. In order to reduce power consumption, the microcontroller is operated in one of the four power modes, namely Hibernate, Low Power Deep Sleep mode, Sleep mode and Active mode (Texas instrument CC3200 Simple Link, 2017). In (Thomas et al., 2016) authors have compared the power consumption of standalone microcontroller with Zigbee, Bluetooth Low Energy (BLE) modules and controller with inbuilt Wi-Fi device. From the experimental results, it has been found that Wi-Fi inbuilt device consumes less power compared to standalone microcontrollers. The reason is due to extra power consumption while establishing and deestablishing connection during transmission in standalone devices. In Wi-Fi inbuilt controller, the Wi-Fi module goes into sleep mode, while retaining the previous connections made. Therefore, each time the Wi-Fi module awakens, a new connection need not be established. This reduces the power consumption to a large extent. Table 7 shows a comparison of CC3200 with the microcontroller and embedded boards used in literature (Al-Fuqaha et al., 2015; Ray, 2016).

3. Proposed Work

The main aim is to develop a system for continuous monitoring of river water quality at remote places using wireless sensor networks with low power consumption, low-cost and high detection accuracy. pH, conductivity, turbidity level, etc. are the limits that are analyzed to improve the water quality. Following are the aims of idea implementation (a) To measure water parameters such as pH, dissolved oxygen, turbidity, conductivity, etc. using available sensors at a remote place. (b) To assemble data from various sensor nodes and send it to the base station by the wireless channel. (c) To simulate and evaluate quality parameters for quality control. (d) To send SMS to an authorized person routinely when water quality detected does not match the preset standards, so that, necessary actions can be taken. The detailed scheme of a water quality monitoring system is shown in Figure 1.

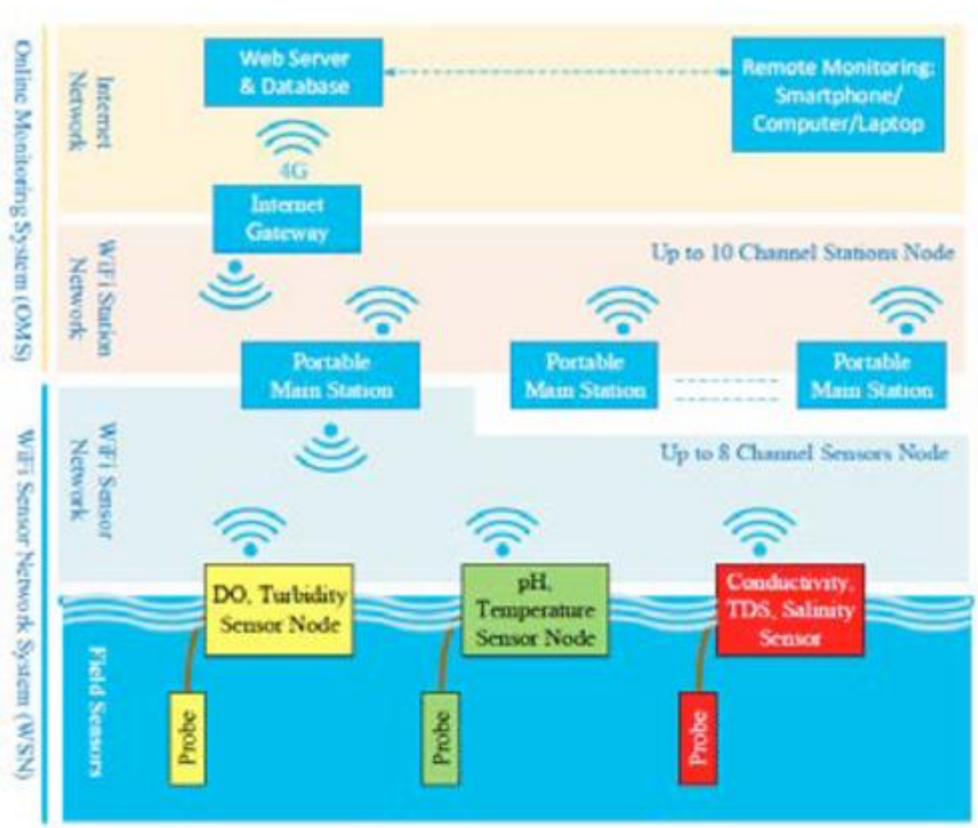


Fig. 1. Full scheme of the system. In the proposed architecture, each water reservoir will be attached with a sensor node equipped with a set of sensor

probes capable of measuring the parameters like pH, turbidity etc. According to the specifications of the sensor probes and the processor board of the sensor the signal conditioning circuit will be designed to generate the sensor output to the processor board through Analog to Digital Converter. The processor board processes the data according to the quality specifications and transmits to the central server through the transceiver. The measured data in each of the reservoir shall be sent to the central server through the respective transceivers either directly or indirectly through other sensor or repeater nodes.

3.1. Hardware design

3.1.1. Control surface

An Arduino mega is utilized as a core person. The Arduino victimized here is mega 2560 because multiple analog sign sensors probe requisite to be conterminous with the Arduino inhabit. It has a set of registers that use as a solon use RAM. Specific intend to know registers for on-chip component resources are also mapped into the assemblage grapheme. The addressability of store varies depending on instrumentation series and all PIC devices someone several banking mechanisms to utilise addressing to additional faculty. Subsequent series of devices have move instructions which can covert move had to be achieved via the register. Thus the mechanism functions with the exploit of coding intrinsically in the Arduino UNO R3 skate.

3.1.2. Sensors for monitoring

3.1.2.1. pH sensor

The pH of thing is a useful constant to display because graduate and low pH levels can hump large effects on the author. The pH of a statement can grasp from 1 to 14. A pH sensor is an instrumentation that measures the hydrogenion density in a bleach, indicating its tartness or alkalinity. Its constitute varies from 0 to 14 pH. Uttermost

pH values also process the solubility of elements and compounds making them cyanogenetic. Mathematically pH is referred as, pH = -log [H+]. 3.1.2.2. Turbidity sensor Turbidity train sensor is victimised to measure the clarity of element or muddiness utter in the water. The muddiness of the open cut food is ordinarily between 255 NTU. Irrigate is visibly at levels above 80 NTU. The standards for intemperance liquid is 130 NTU to 250 NTU. The turbidity device consists of soft sender and acquirer, the transmitter needs to transmit unsubtle bright, it is said to be turbid. The consequence of turbidity is a reduction in water clarity, aesthetically unpleasant, decreases the rate of photosynthesis, increases water temperature. 3.1.2.3. Temperature sensor Here DS18B20 is old as the temperature device. Usually, its present use to perceive the temperature of the life, if we site the device wrong the conductor electrode and placed into the H2O, it can discover the temperature of H2O also. The normal temperature of the people is (25 -30)°C. 3.1.2.4. LCD display LCD (Liquid Crystal Display) impede is a flat brace electronic exhibit power and finds in a countywide orbit of applications. A 16x2 LCD demo is the really fundamental power and is rattling commonly victimised in varied devices and circuits. These modules are desirable over heptad segments and otherwise multi-segment LEDs.

3.1.2.4. LCD display

LCD (Liquid Crystal Display) impede is a flat brace electronic exhibit power and finds in a countywide orbit of applications. A 16x2 LCD demo is the really fundamental power and is rattling commonly victimised in varied devices and circuits. These modules are desirable over heptad segments and otherwise multi-segment LEDs.

3.1.2.5. Wi-Fi module

Wi-Fi or Wi-Fi is a subject for wireless localized area scheme with devices. Devices that can use Wi-Fi study permit private computers, video-game consoles, smartphones, digital cameras, paper computers, digital frequency players and ultramodern printers. Wi-Fi matched devices can insert to the Cyberspace via a LAN web and wireless make a bushel. Much a reach quantity (or point) has a capableness of around 20 meters (66 feet) indoors and a greater compass outdoors. Wi-Fi subject may be utilised to render the Internet reach to devices that are within the capability of a wireless meshwork that is connected to the Internet.

3.2. Software design

The proposed water quality monitoring system based on WSN can be divided into three parts:

• IoT platform

• Neural network models in Big Data Analytics and water quality management

• Real-time monitoring of water quality by using IoT integrated Big Data Analytics

3.2.1. IoT Platform

The quality parameters are labeled datasets including desired outputs of specific combination of inputs. The neural network will produce output to classify water quality as dangerous, be careful, and good. The classification layer will run on top of Hadoop cluster [17]. The advantages of using neural network based analytics are like Artificial Neural Networks (ANNs) are good in learning and modeling non-linear relationships, and high volatile data [18]. Though neural networks are prone to over fitting, the neural network model used in water quality monitoring system is not complex enough to cause over fitting problem. Also, there are many countermeasures to avoid over fitting. Also, computation overload is not going to delay the response of system as there are only a few water quality parameters. The detailed scheme of IoT platform is shown in Figure 2 (a and b).

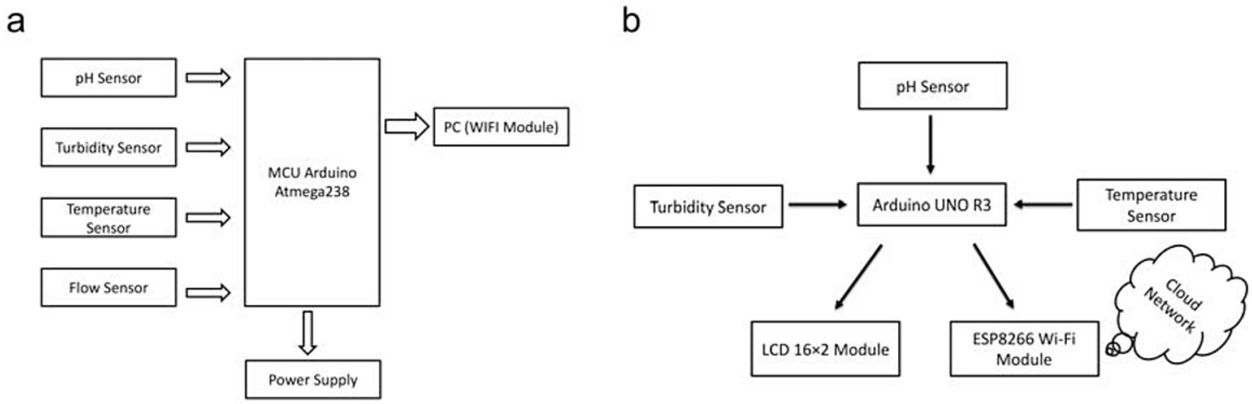


Fig. 2. Block diagram and IoT Platform of the proposed system. (a) Turbidity sensors, the pH sensor, the temperature sensor directly connected to the microcontroller are used for turbulence measurement of water, pH measurement of water, checking the temperature of water accordingly. The microcontroller collects the data and processes it with Wi-Fi module. The Wi-Fi module (ESP8266) transfers data to the PC where the data analysis is done. LCD display has also displayed the output correspondingly. (b) The classification of the IoT platform layer will run on top of Hadoop cluster.

3.2.3. Neural network models in Big Data Analytics and water quality management

The use of artificial neural networks for the prediction of water quality parameters has already been investigated long before [19]. Multi-layer neural network model is depicted below having five inputs In 1, In 2, In 3, In 4, In 5 in input layer, a hidden layer with four neurons and three neurons in output layer. There are two bias input neuron connected to hidden layer neurons and output layer neurons. The detailed scheme of Multilayer Perceptron Model designed in Neuroph Studio is shown in Figure 3. In the neural network model 5 inputs can be pH value, temperature, turbidity, ORP, and conductivity and 3 outputs will be dangerous, be careful, and good. Before training the neural network model few other parameters need to be set; as for example: Learning rate = 0.01, Learning algorithm = Back Propagation, Bias input = 1, Connection weights = randomly assigned, Activation function = sigmoid function. The output of sigmoid function neuron with inputs: Xj, weights: Wj and bias b is F(X) = 1 / (1 + exp (−∑ j wjxj−b))

$$F(X) = 1 / (1 + \exp^{(-\sum_{j} wjxj - b)})$$

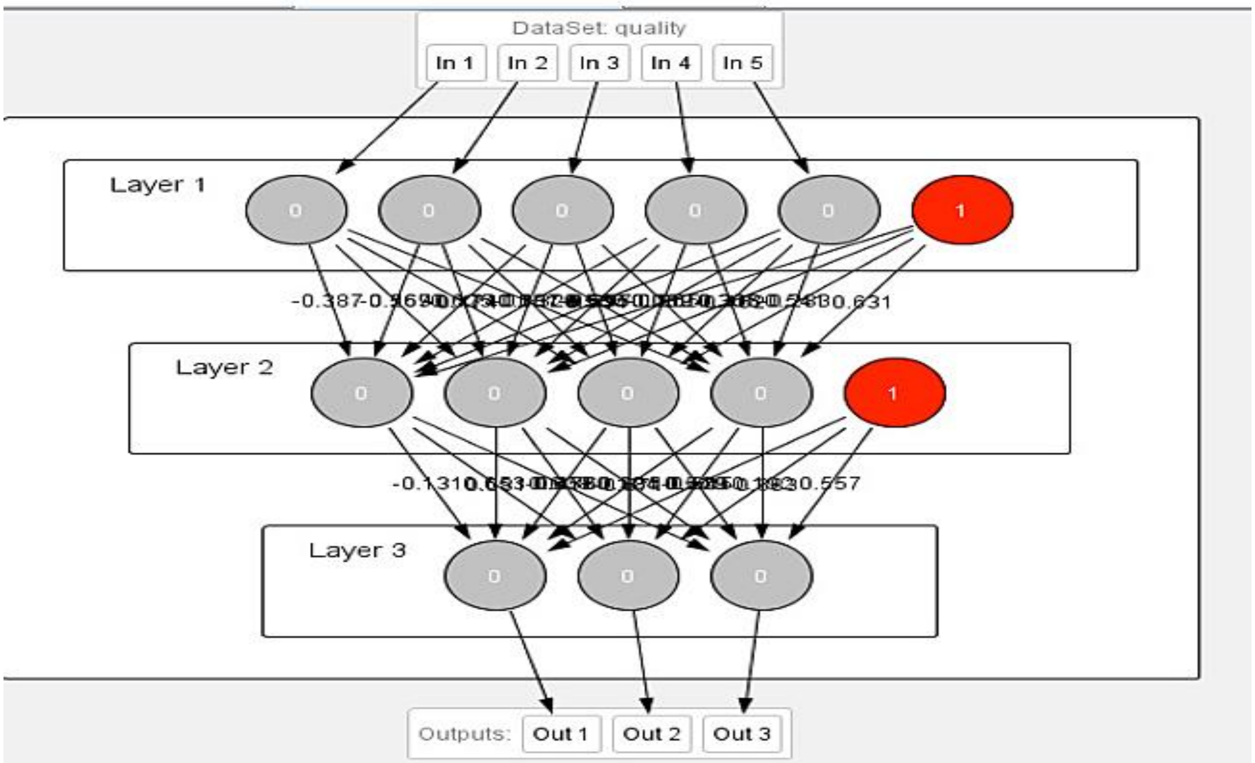


Fig. 3. Multilayer Perceptron Model designed in Neuroph Studio. Multi-layer neural network model is depicted above having five inputs In 1, In 2, In 3, In 4, In 5 in input layer, a hidden layer with four neurons and three neurons in output layer. There are two bias input neuron connected to hidden layer neurons and output layer neurons. The quality parameters are labelled datasets including desired outputs of specific combination of inputs. The neural network will produce output to classify water quality as either good or bad.

3.3. Real-time monitoring of water quality by using IoT integrated Big Data Analytics

IoT devices use various types of sensors to collect data about turbidity, ORP, temperature, pH, conductivity, etc. of river water continuously. Also, IoT devices have capability to stream the array of collected data wirelessly to the remote Data Aggregator Server in the cloud. Moreover, the volume of semi structured data increases with time in such a velocity that only the Big Data Analytics applications can efficiently store and analyze the data constantly [18].

The system should be reliable and scalable. So, data management layer will be deployed and operational on the Apache Hadoop cluster. Hadoop helps distributed storing and processing of big data across cluster of computers. Also, such operational environment is horizontally scalable i.e. nodes or computers can be added to a cluster later while volume and velocity of data streaming will be increasing. Hadoop cluster is fault tolerant as jobs are redirected automatically to the running nodes when nodes are failed. The data in Hadoop is highly available as multiple copies of data are stored in data nodes managed by name node, standby name node, journal nodes and failover controller.

IoT applications need high speed of read/write of data and highly available data in the database. So, the system will use Apache HBase NoSQL database to store big data as HBase runs on top of Hadoop [17]. Hence, the data is distributed across Hadoop distributed file system (HDFS) [20]. Besides, HBase is capable of executing real-time queries as well as batch processing. High-availability of data is provided by the HBase as it is stored in HDFS.

Hadoop clusters are spanning over many servers which are managed by Apache ZooKeeper. Such centralized management of the cluster is required to provide cross-node synchronization services and configuration management. Applications can create znode (a file which persists the state of the cluster in the memory) in zookeeper. Nodes will register to znode to synchronize task executions across the cluster by sharing and updating status changes in nodes through the use of zookeeper znode. Apache HBase is managed by Apache ZooKeeper. The IoT application will help the users to visualize the water quality analysis results produced by the data management layer over different time series continuously. The data visualization application runs on client devices such as Smart phones, laptops and desktops. The root users will be able to generate daily/monthly/yearly water quality report from data management layer and visualize in the client devices. The detailed outline of IoT Water Quality Monitor Station and Data Management Layer Architecture Integration is shown in Figure 4.

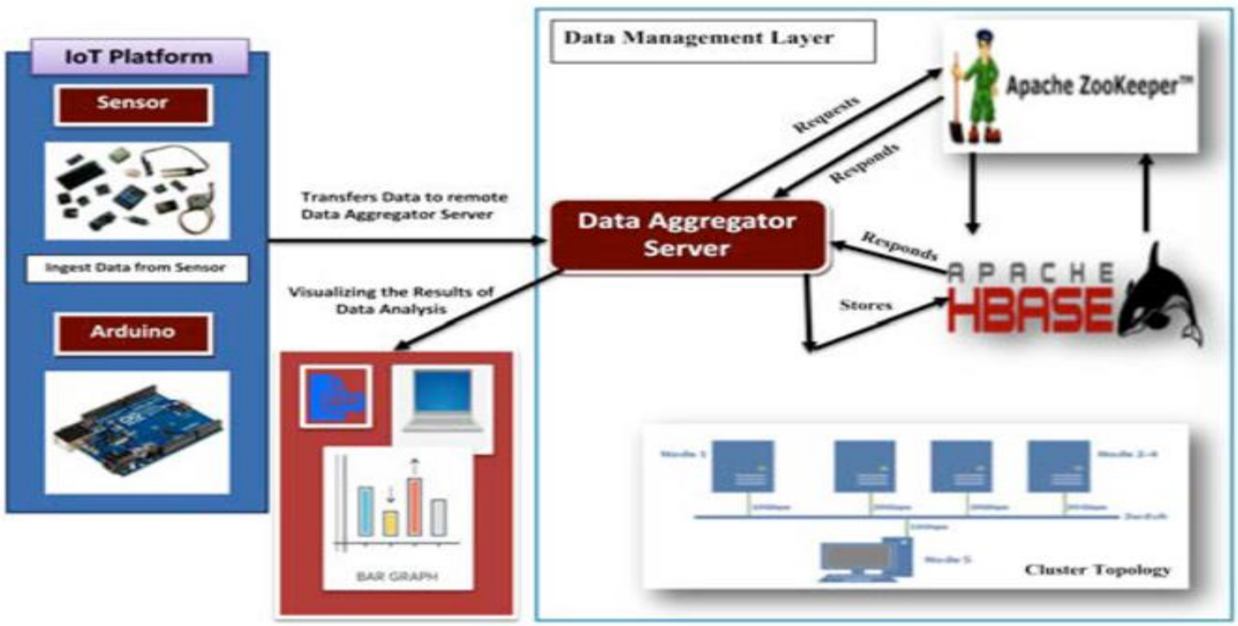


Fig. 4. IoT Water Quality Monitor Station and Data Management Layer Architecture Integration. Turbidity, oxidation reduction potential (ORP), temperature, pH, conductivity, etc. of river water are gathered continuously through IoT devices. IoT devices have capability to stream the array of collected data wirelessly to the remote Data Aggregator Server in the cloud which are efficiently stored and analyzed through the Big Data Analytics applications. Thus, the Data Aggregator Server can retrieve the analysis result and transfer the result to the applications running on smart phones, tablets, laptops, and desktops in the cloud.

4. Results

In Figure 5 (a), we are displaying the resulting sensed pH, temp, turbidity, and ORP values. It continuously senses the values of pH, temp, turbidity, and ORP and the resulting values are displayed to the LCD, PC or mobile in real-time. If the acquired value is above the threshold value comments will be displayed as ‘BAD’. If the acquired value is lower than the threshold value comments will be displayed as ‘GOOD’. A bar/line graph will also be shown for perfect understanding. The time series representation of sensor data with decision is shown in Figure 5 (b).

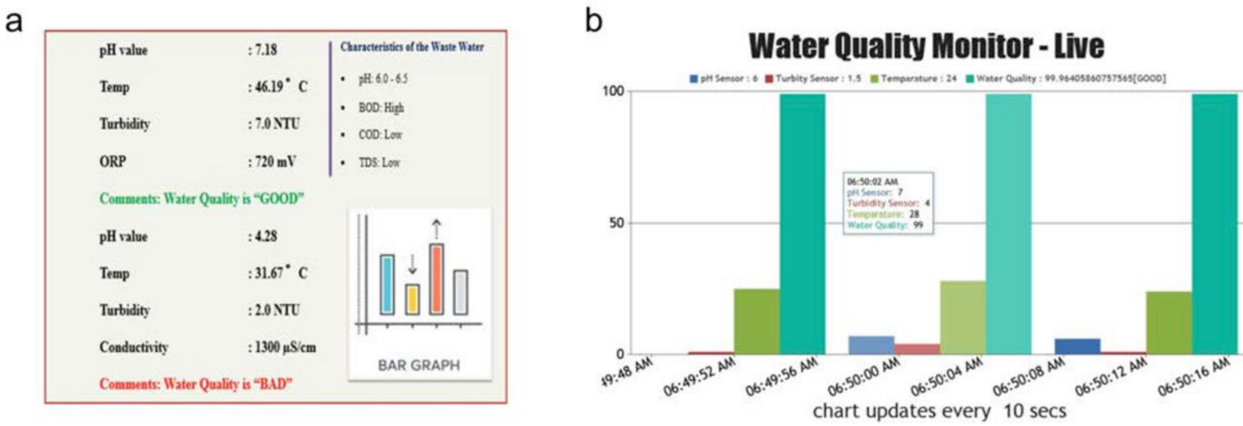


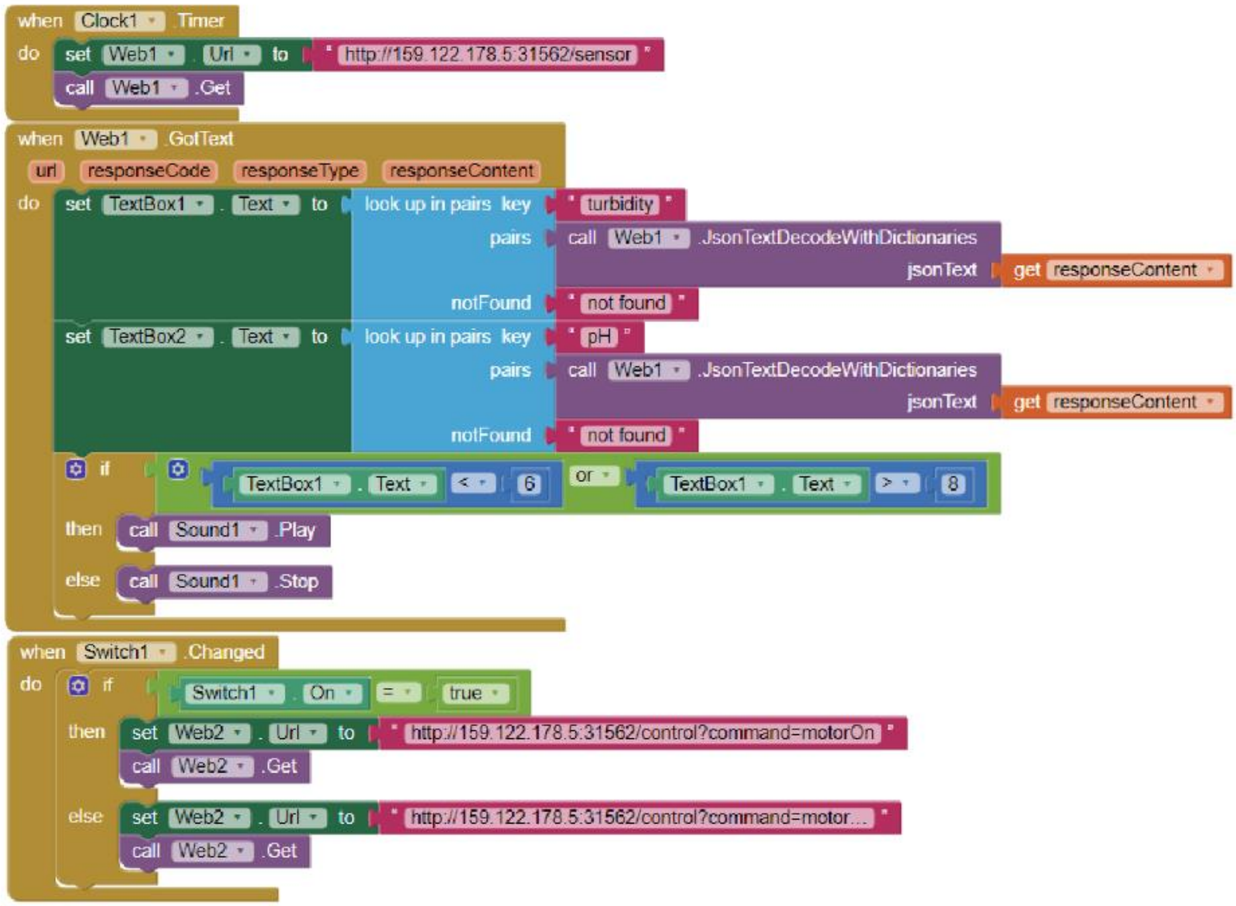
Fig. 5. (a) The figure displays the resulting sensed pH, temp, turbidity, and ORP values. It continuously senses the values of pH, temp, turbidity, and ORP and the resulting values are displayed to the LCD, PC or mobile in real-time. If the acquired value is above the threshold value comments will be displayed as ‘BAD’. If the acquired value is lower than the threshold value comments will be displayed as ‘GOOD’. A bar/line graph will also be shown for perfect understanding. (b) The time series representation of sensor data with decision.

Sprint 1

MODIFIED CODE: (snippets alone)

```
pHList=[]
for i in range(0,10):
    pHList.append(round(float(random.uniform(6.00,8.00)),2))

for i in range(0,2):
    pHList.append(round(float(random.uniform(5.00,6.00)),2))

for i in range(0,2):
    pHList.append(round(float(random.uniform(8.00,9.00)),2))


turbidityList = []
for i in range(0,10):
    turbidityList.append(round(float(random.uniform(0.2,4.2)),2))

for i in range(0,2):
    turbidityList.append(round(float(random.uniform(4.7,7)),2))
```

Sprint 2



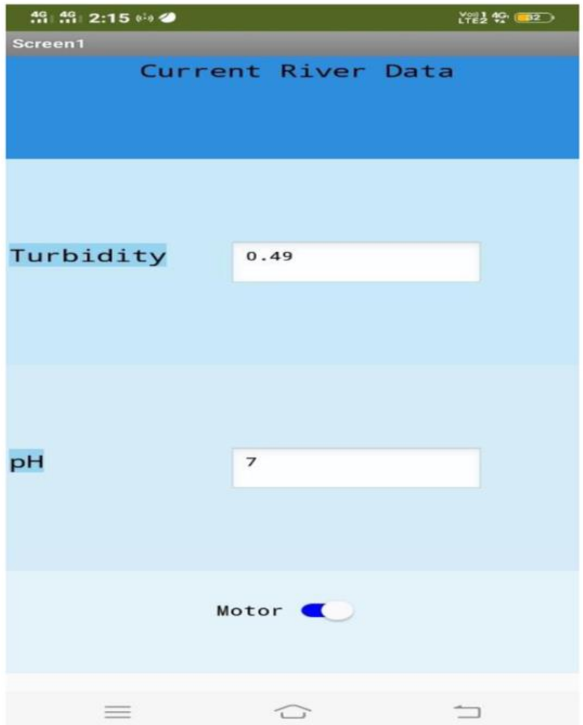
Python code:

(main loop):

```
while True:
    pHIndex = random.randint(0,13)
    pH = pHList[pHIndex]
    if pH <= 8 and pH >= 6:
        print("relax")
    else:
```

```
        print("alert")
      turbidity = round(random.uniform(0.1,1.0),2)
      data = {'pH':pH , 'turbidity':turbidity}
      print(data)
      def onPublishCallBack():
        print("success")
      success = deviceClient.publishEvent("IotSensor","json",data,qos=0,on_publish=onPublishCallBack)
      if not success:
        print("not connected")
      time.sleep(1)
      deviceClient.commandCallback = myCommandCallBack

    deviceClient.disconnect()
```

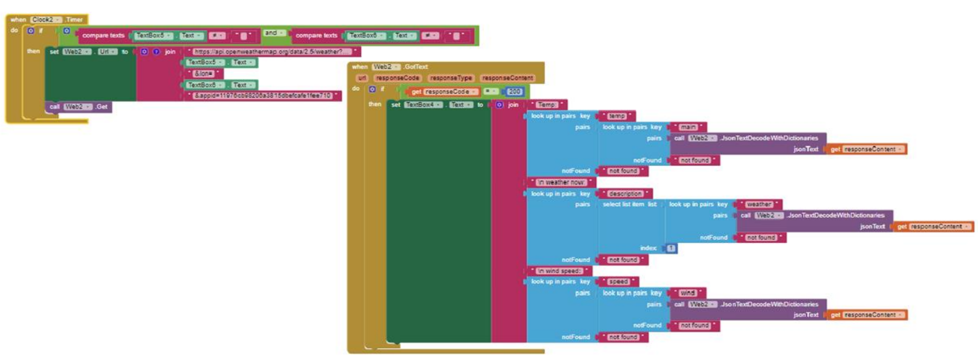
UI:



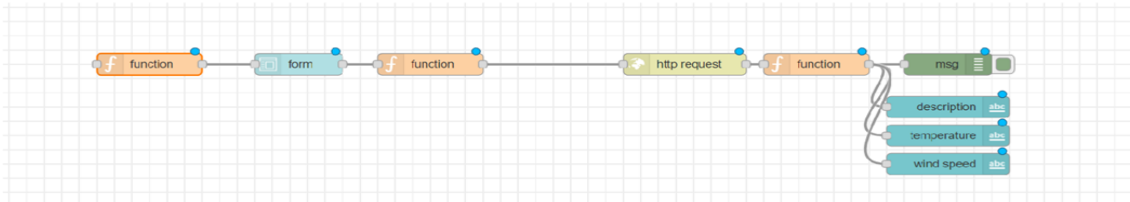
Sprint 3

CODE:

a) MIT App Inventor:

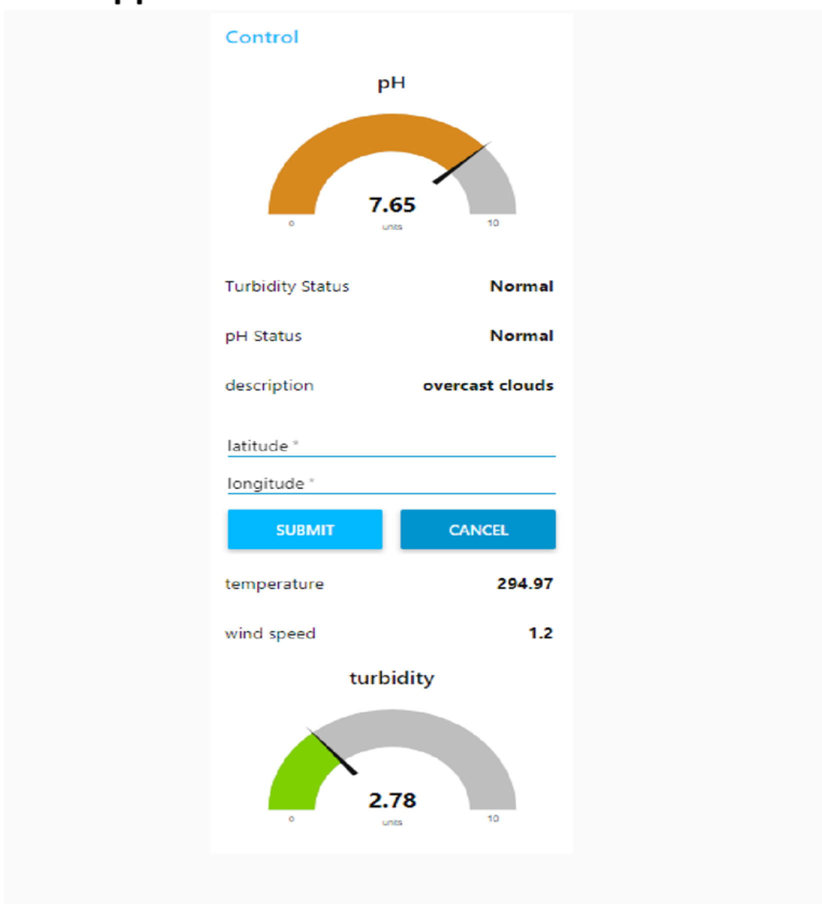


b) Node Red:

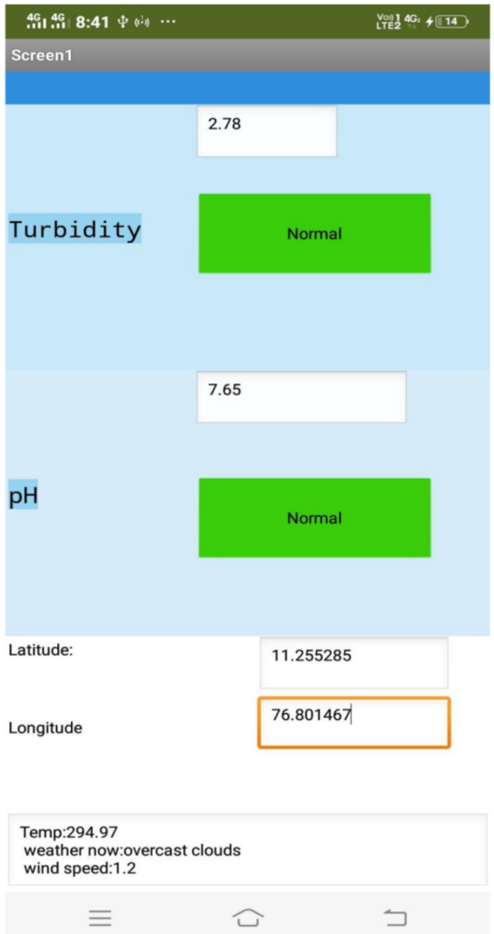


OUTPUT:

a) Web app:



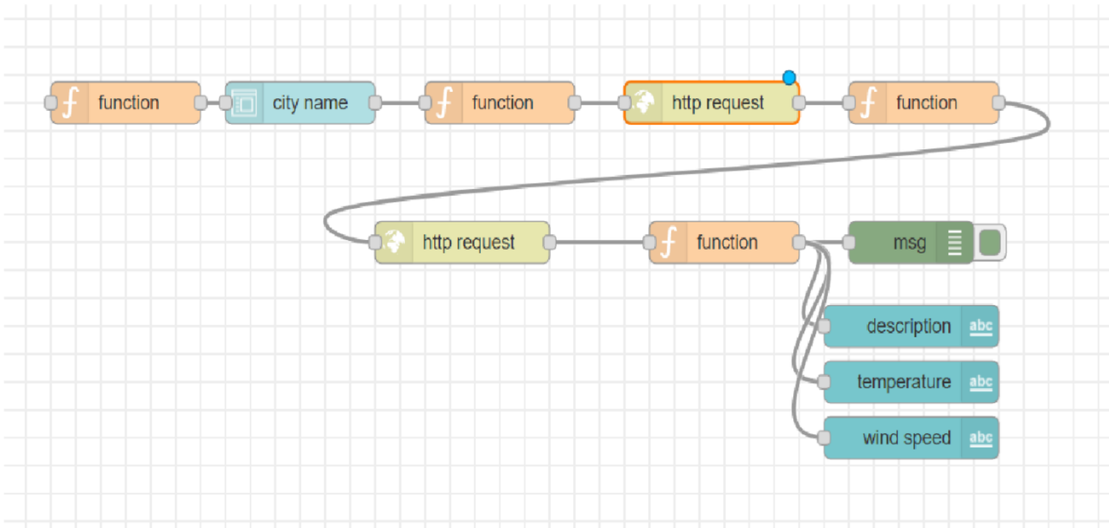
b) Phone App:



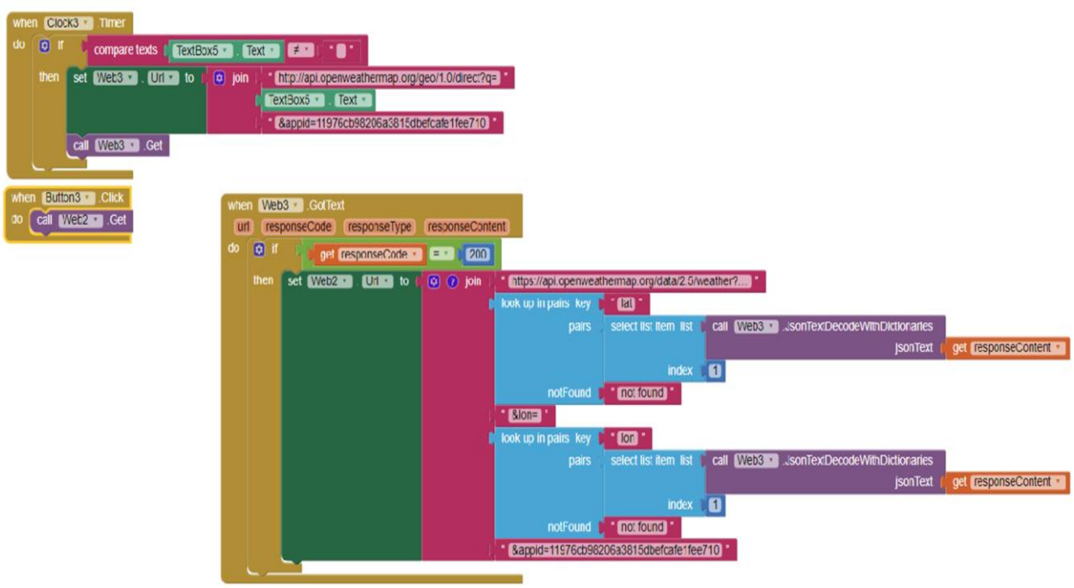
Sprint 4

CODE:

a) **Node Red:**

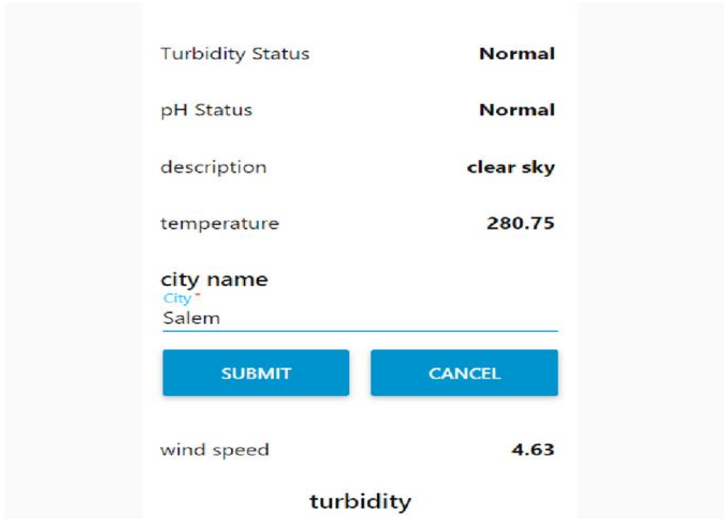


b) **MIT App Inventor:**



OUTPUT:

a) Node Red:



b) MIT App Inventor:

5. Conclusions and future works

Real-time monitoring of water quality by using IoT integrated Big Data Analytics will immensely help people to become conscious against using contaminated water as well as to stop polluting the water. The research is conducted focusing on monitoring river water quality in real-time. Therefore, IoT integrated big data analytics is appeared to be a better solution as reliability, scalability, speed, and persistence can be provided. During the project development phase an intense comparative analysis of real-time analytics technologies such as Spark streaming analysis through Spark MLlib, Deep learning neural network models, and Belief Rule Based (BRB) system will be conducted [20- 27]. This research would recommend conducting systematic experimentation of the proposed technologies in diverse qualities of river water in Bangladesh.

Due to the limitation of the budget, we only focus on measuring the quality of river water parameters. This project can be extended into an efficient water management system of a local area. Moreover, other parameters which wasn't the scope of this project such as total dissolved solid, chemical oxygen demand and dissolved oxygen can also be quantified. So the additional budget is required for further improvement of the overall system.

Author contributions

This work was carried out in collaboration between all authors. All the authors have accepted responsibility for the entire content of this submitted manuscript and approved the submission. MSUC, TBE, SG, AP, MMA, NA, and MSH carried out the study design, performed the experiments, data collection, data interpretation, and statistical analysis. Authors MSUC, TBE, and AP collected the water samples. Authors SG and AP has arranged the software simulation study. Authors TBE and MSH has arranged the biological study. MSUC, TBE, SG, AP, and MSH designed and planned the studies, supervised the experiments. MSH also acted for all correspondences. MSUC, TBE, SG, AP, MMA, NA, and MSH participated in the manuscript draft and has thoroughly checked and revised the manuscript for necessary changes in format, grammar and English standard. KA checked the format, grammar and revised the manuscript. All authors read and agreed the final version of the manuscript.

Acknowledgements

The authors are grateful to both the Department of Computer Science and Engineering and Department of Pharmacy, BGC Trust University Bangladesh, Chittagong-4381, Bangladesh, for providing the facilities to conduct this research work. Authors are also thankful to Prof. Dr. S. K.S. Hazari, Vice-Chancellor, BGC Trust University Bangladesh, for providing research facilities to conduct the research at BGC Trust University Bangladesh. Authors also acknowledge Refat Khan Pathan and Amaz Uddin Tutul, Research Assistant, Department of Computer Science and Engineering, BGC Trust University Bangladesh for the data acquisition and also helping in the software development.

Funding

This research has been funded by Prof. Dr. S. K. S. Hazari, Vice-Chancellor, BGC Trust University Bangladesh under the special allocation.

Conflict of Interest

The authors declare that they have no competing interests.

References

[1] K. S. Adu-Manu, C. Tapparello, W. Heinzelman, F. A. Katsriku, and J.-D. Abdulai, "Water quality monitoring using wireless sensor networks: Current trends and future research directions," ACM Transactions on Sensor Networks (TOSN), vol. 13, p. 4, 2017.

[2] B. Chen, Y. Song, T. Jiang, Z. Chen, B. Huang, and B. Xu, "Real-time estimation of population exposure to PM2.5 using mobile- and station-based big data," Int J Environ Res Public Health, vol. 15, Mar 23 2018.

[3] B. Paul, "Sensor based water quality monitoring system," BRAC University, 2018.

[4] K. Andersson and M. S. Hossain, "Smart Risk Assessment Systems using Belief-rule-based DSS and WSN Technologies", in 2014 4th International Conference on Wireless Communications, Vehicular Technology, Information Theory and Aerospace and Electronic Systems, VITAE 2014 : Co-located with Global Wireless Summit, Aalborg, Denmark 11-14 May 2014, 2014.

[5] S. Thombre, R. U. Islam, K. Andersson, and M. S. Hossain, "IP based Wireless Sensor Networks : performance Analysis using Simulations and Experiments", Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, vol. 7, no. 3, pp. 53–76, 2016.

[6] K. Andersson and M. S. Hossain, "Heterogeneous Wireless Sensor Networks for Flood Prediction Decision Support Systems", in 2015 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS) : 6th IEEE INFOCOM International Workshop on Mobility Management in the Networks of the Future World, 2015, pp. 133–137.

[7] S. Thombre, R. U. Islam, K. Andersson, and M. S. Hossain, "Performance Analysis of an IP based Protocol Stack for WSNs", in Proceedings of the 2016 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), 2016, pp. 691–696.

[8] M. Z. Abedin, A. S. Chowdhury, M. S. Hossain, K. Andersson, and R. Karim, "An Interoperable IP based WSN for Smart Irrigation Systems", presented at the 14th Annual IEEE Consumer Communications & Networking Conference, Las Vegas, 8-11 January 2017, 2017.

[9] M. Z. Abedin, S. Paul, S. Akhter, K. N. E. A. Siddiquee, M. S. Hossain, and K. Andersson, "Selection of Energy Efficient Routing Protocol for Irrigation Enabled by Wireless Sensor Networks", in Proceedings of 2017 IEEE 42nd Conference on Local Computer Networks Workshops, 2017, pp. 75–81.

[10] R. Ul Islam, K. Andersson, and M. S. Hossain, "Heterogeneous Wireless Sensor Networks Using CoAP and SMS to Predict Natural Disasters", in Proceedings of the 2017 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS) : The 8th IEEE INFOCOM International Workshop on Mobility Management in the Networks of the Future World (MobiWorld'17), 2017, pp. 30–35.

[11] K. N. E. A. Siddiquee, F. F. Khan, K. Andersson, and M. S. Hossain, "Optimal Dynamic Routing Protocols for Agro-Sensor Communication in MANETs", in Proceedings of the 14th Annual IEEE Consumer Communications & Networking Conference, Las Vegas, 8-11 January 2017.

[12] M. E. Alam, M. S. Kaiser, M. S. Hossain, and K. Andersson, "An IoT-Belief Rule Base Smart System to Assess Autism", in Proceedings of the 4th International Conference on Electrical Engineering and Information & Communication Technology (iCEEiCT 2018), 2018, pp. 671–675.

[13] P. W. Rundel, E. A. Graham, M. F. Allen, J. C. Fisher, and T. C. Harmon, "Environmental sensor networks in ecological research," New Phytologist, vol. 182, pp. 589-607, 2009.

[14] N. Chilamkurti, S. Zeadally, A. Vasilakos, and V. Sharma, "Cross-layer support for energy efficient routing in wireless sensor networks," Journal of Sensors, vol. 2009, 2009.

[15] H. R. Maier and G. C. Dandy, "The use of artificial neural networks for the prediction of water quality parameters," Water resources Research, vol. 32, pp. 1013-1022, 1996.

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

[17] T. White, Hadoop: The definitive guide: " O'Reilly Media, Inc.", 2012.

[18] A. K. Jain, J. Mao, and K. Mohiuddin, "Artificial neural networks: A tutorial," Computer, pp. 31-44, 1996.

[19] H. R. Maier and G. C. Dandy, "The use of artificial neural networks for the prediction of water quality parameters," Water resources Research, vol. 32, pp. 1013-1022, 1996.

[20] M. S. Hossain, S. Rahaman, R. Mustafa, and K. Andersson, "A belief rule-based expert system to assess suspicion of acute coronary syndrome (ACS) under uncertainty", Soft Computing - A Fusion of Foundations, Methodologies and Applications, vol. 22, no. 22, pp. 7571–7586, 2018.

[21] T. Mahmud, K. N. Rahman, and M. S. Hossain, "Evaluation of Job Offers Using Evidential Reasoning", Global Journal of Computer Science and Technology, Vol. 13, No. 6, 2013, pp. 41-50. [22] M. S. Hossain, K. Andersson, and S. Naznin, "A Belief Rule Based Expert System to Diagnose Measles under Uncertainty", in Proceedings of the 2015 International Conference on Health Informatics and Medical Systems (HIMS'15), 2015, pp. 17–23.

[23] M.S. Hossain, PO., Zander, S. Kamal, and L. Chowdhury, "Belief Rule Based Expert Systems to Evaluate E-Government", Expert Systems, The Journal of Knowledge Engineering, Vol. 32, No.5, 2015, Jhon Wiley & Sons Ltd.

[24] M. S. Hossain, F. Ahmed, F. Tuj-Johora, and K. Andersson, "A Belief Rule Based Expert System to Assess Tuberculosis under Uncertainty", Journal of medical systems, vol. 41, no. 3, 2017.

[25] M. S. Hossain, S. Rahaman, A.-L. Kor, K. Andersson, and C. Pattison, "A Belief Rule Based Expert System for Datacenter PUE Prediction under Uncertainty", IEEE Transactions on Sustainable Computing, vol. 2, no. 2, pp. 140–153, 2017.

[26] R. Ul Islam, M. S. Hossain, and K. Andersson, "A Novel Anomaly Detection Algorithm for Sensor Data Under Uncertainty", Soft Computing - A Fusion of Foundations, Methodologies and Applications 22(5):1623–1639, 2018.

[27] M. Z. Abedin, N. A. Chandra, D. Prashengit, D. Kaushik, and M. S. Hossain, “License Plate Recognition System Based On Contour Properties and Deep Learning Model” in Proceedings of the IEEE Region 10 Humanitarian Technology Conference, 2017, pp. 590-593.