Ideation Phase
Ideation

DATE	21 October 2022
TEAM ID	PNT2022TMID23696
PROJECT NAME	Real Time River Water Quality Monitoring and Control System
MAXIMUM MARKS	4 Marks

Ideation phase:

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 like pH, dissolved oxygen, turbidity, conductivity,etc.
b) To assemble data from various sensor nodes and send it to the base station by the
c) wireless channel.
d) To simulate and evaluate quality parameters for quality control.
e) To send SMS to an authorized person routinely when water quality detected does not
f) match the present standards, so that, necessary actions can be taken.

Control surface:

An Arduino mega is utilized. The Arduino board used here is mega 2560. It has a set of registers that use as a solon use RAM.The Arduino Mega 2560 is a microcontroller board based on the ATmega2560. It has 54 digital input/output pins (of which 14 can be used as PWM outputs), 16 analog inputs, 4 UARTs (hardware serial ports), a 16 MHz crystal oscillator, a USB connection, a power jack, an ICSP header, and a reset button. It contains everything needed to support the microcontroller; simply connect it to a computer with a USB cable or power it with a AC-to-DC adapter or battery to get started. The Mega is compatible with most shields designed for the Arduino Duemilanove or Diecimila.

pH sensor:

A pH sensor is one of the most essential tools that's typically used for water measurements. This type of sensor is able to measure the amount of alkalinity and acidity in water and other solutions. When used correctly, pH sensors are able to ensure the safety and quality of a product and the processes that occur within a wastewater or manufacturing plant. In most cases, the standard pH scale is represented by a value that can range from 0-14. When a substance has a pH value of seven, this is considered to be *neutral*. Substances with a pH value above seven represent higher amounts of alkalinity whereas substances with a pH value that's lower than seven are believed to be more acidic.

Mathematically pH is referred as,
pH = -log [H+].

Turbidity sensor:

Turbidity is the measurement of water clarity (i.e., transparency). Suspended particles – such as silt, algae, plankton, and sewage – can cause water to appear cloudy or murky. These particles scatter and absorb light rays rather than allowing light to be transmitted straight through the water. A higher turbidity reading represents cloudier and 'thicker' water with more particles throughout. When water is clear, it has low turbidity levels.
Turbidity sensors measure the amount of light that is scattered by the suspended solids in water. As the amount of total suspended solids (TSS) in water increases, the water's turbidity level (and cloudiness or haziness) increases.

Temperature sensor:

The DS18B20 is a 1-wire programmable Temperature sensor from maxim integrated. It is widely used to measure temperature in hard environments like in chemical solutions, mines or soil etc. The constriction of the sensor is rugged and also can be purchased with a waterproof option making the mounting process easy. It can measure a wide range of temperature from -55°C to +125° with a decent accuracy of ±5°C. Each sensor has a unique address and requires only one pin of the MCU to transfer data so it a very good choice for measuring temperature at multiple points without compromising much of your digital pins on the microcontroller.

LCD display:

LCD is a flat display technology, stands for "Liquid Crystal Display," which is generally used in computer monitors, instrument panels, cell phones, digital cameras, TVs, laptops, tablets, and calculators. It is a thin display device that offers support for large resolutions and better picture quality. A 16x2 LCD demo is the really fundamental power and is used here.

Wi-Fi module:

Wi-Fi is a subject for wireless localized area scheme with devices. Devices that canuse Wi-Fi are 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. Wi-Fi signals will usually reach about 150 feet or over 45 meters for a 2.4Ghz frequency. Using a 5Ghz frequency, we will get about 50 feet or about 15 meters of reach.

Software design:

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

a) IoT platform
b) Neural network models in Big Data Analytics and water quality management
c) Real-time monitoring of water quality by using IoT integrated Big Data Analytics

IoT Platform:

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 dangerous, be careful, and good. The classification layer will run on top of Hadoop cluster. The advantages of using neural network based analytics are like Artificial Neural Networks (ANNs) are good in learning and modelling non-linear relationships, and high volatile data.

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.

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 . 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 neurons connected
to hidden layer neurons and output layer neurons.
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 :

$$F(X) = 1 / (1 + \exp(- \Sigma jwjxj - b))$$

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 analyse the data constantly . 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 HB se NoSQL database to store big data as HBase runs on top of Hadoop . Hence, the data is distributed across Hadoop distributed file system (HDFS) . 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 Zoo Keeper. Such centralized management of the cluster is required to provide cross-node synchronization services and configuration management. Applications can create z node (a file which persists the state of the cluster in the memory) in zookeeper. Nodes will register to z node to synchronize task executions across the cluster by sharing and updating status changes in nodes through the use of

zookeeper z node. Apache HBase is managed by Apache Zoo Keeper. 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.