**Abstract:**

To address the challenges of high labor costs, lengthy detection periods, and insufficient information in current environmental monitoring systems, this paper proposes a water quality monitoring system based on LoRa, Cellular Network, and Machine Learning (ML) with low-power Internet of Things (IoT) technology. The system features data storage, a web application user interface, long-distance data transmission, dynamic monitoring, AI prediction, data visualization, and pollution alarms for the distributed deployment of multisensory node information (pH, turbidity, total dissolved solids, and water temperature). The system utilizes an Arduino Nano equipped with various water quality sensors to collect real-time water quality parameters. The collected data is then packaged and sent to an ESP32, which collaborates with a GSM SIM800C module to function as a remote gateway using LoRa technology. The data is stored on an SD card. The gateway bridges the LoRa link to an IP link, forwarding the water quality information to a Firebase Cloud server. Finally, end-users can monitor and control water quality through a web/app platform. In the experiments after testing on 3 different lakes in National Polytechnic Institute of Cambodia results show that the system has a good performance in terms of real time data and acquisition accuracy, data transmission reliability, Pollution alarm success rate. The average relative errors of water temperature, pH, turbidity, and conductivity are 0.31%, 0.28%, 3.96%, and 0.71%, respectively. In addition, the signal reception strength of the system within 2 km is better than -81 dBm, and the average packet loss rate is only 94%. In short, the system’s high accuracy, high reliability, long-distance characteristics meet the needs of large area water quality monitoring and the prediction having more accurately than we expect as our data have been train many times so the accuracy are 63%.

**Introduction**

Water is the natural resource for the survival of mankind and is of great significance to human production and life. In recent years, with the vigorous development of mankind, domestic sewage, production wastewater and various wastes discharged from agricultural production are directly discharged into water bodies without treatment, which causes serious pollution of rivers, lakes and groundwater, further leads to serious deterioration of the water quality in the area, affects the normal life of residents and causes ecological unbalance. Therefore, The quality of drinking water is important to human health and to provide a safe drinking water supply is one of the main objectives of Cambodian National Policy. Cambodia is located in Southeast Asia between latitudes 10° and 15° N. and longitudes 102° and 108° E. The country covers an area of 181,035 km². Cambodia is bordered by Vietnam in the east and southeast, the Lao PDR in the north and by Thailand in the north and northwest. To the southwest the country has a seacoast on the Gulf of Thailand. In Cambodia, both surface water and groundwater are used for drinking water. The Mekong River and the Tonle Sap Lake are the predominant sources of surface water, with the Mekong serving the east and the Great Lake serving the more westerly populations. The river system provides abundant and good quality drinking water. Applying the WHO standards, these resources require only basic treatment including disinfection. Provincial towns generally have access to surface water from the river systems in unlimited quantities. Although Cambodia has abundant water resources but the accelerating pace of industrial development and population growth in recent decades have affect the quality of water. Since Cambodia is local in Mekong River, In the recognizing that sustainable development of water resources of the LMB will not be possible without effective management of water quality, the MRC Member Countries (MCs) agreed to establish a Water Quality Monitoring Network (WQMN) with the specific objective of detecting changes in the Mekong River water quality and ensuring that preventive and remedial actions are taken if any changes are detected. The routine monitoring and reporting of water quality are the main functions of the WQMN, which was established in 1985 with a funding support from the Swedish International Development Agency (Sida). In 2018 Forty-eight (48) stations were monitored by the WQMN in 2018. There 11 stations in Lao PDR, 8 Station in Thailand, 19 Station in Cambodia, 10 Station in Viet Nam. For consistency, the MCs have agreed to carry out the sampling and monitoring of water quality on a monthly basis between the 13th and 18th day of each month. In order to monitoring the quality MCR…………………………………………………………………………………………………………………………………………………………..

**Architecture Design**

The NPIC lake water quality monitoring system proposed by this research consists of four parts: perception layer, transmission layer, Machine Learning, platform layer, and application layer. The system mainly realizes the functions for distributed collection of water quality data, node positioning, remote transmission, data storage, remote monitoring and AI Prediction. The system architecture diagram is shown in Figure 1.

The lake water quality monitoring system proposed by this research consists of four parts: perception layer, transmission layer, platform layer, and application layer. The system mainly realizes the functions for distributed collection of water quality data, node positioning, remote transmission, data storage, and remote monitoring. The system architec- ture diagram is shown in Figure 1.

The water quality monitoring node in this system is based on LoRa technology. The node is distributed in the target water area and consists of a control unit, a water temperature-pH composite sensor, a turbidity sensor, a con- ductivity sensor, a power management module, and a LoRa Radio Frequency (RF) transceiver module. On the one hand, the LoRa node collects various water quality parameters such as water temperature, pH, turbidity, and conductivity by sensors. Finally, the information is packaged and sent to the transport layer by the LoRa communication module. In the transport layer, in order to cope with high PLR (packet loss rate) that may be caused by the data access of large-scale nodes, this research has extended eight RFM92 baseband chips for the RFM92 baseband chip of the LoRa gateway. In this way, the symmetry of uplink and downlink eight channels is realized, and a reliable trans- mission link is provided for user data. The third layer is the platform layer, which is responsible for aggregating terminal data forwarded by the Gateway. And according to the diﬀerent data types, it is stored in the Firebase database in an orderly manner and provides support for the monitoring application system to realize specific business functions. The application layer is the fourth layer. The monitoring system completes data analysis, query, visualization, local storage, pollution alarm, AI Prediction and other functions by calling the data processing interface provided by the Web/App platform.

The system has a clear hierarchy from bottom to top. The terminal node of the perception layer obtains detailed data. The transport layer puts forward countermeasures in the face of the large-scale data access problem of distributed nodes. The platform layer provides reliable support for user applications. The monitoring system at the application layer is fully functional.

**Method**

**Node Design**

The water quality monitoring node is located in the sensing layer of monitoring system and is distributed in the monitoring target water area. Each node has its unique ID number and communicates with the gateway by diﬀerent channels. It can be divided into five parts for design: main control board design, LoRa RF unit selection, power management module design, and data frame format design. Among them, the design of the main control board and LoRa RF unit is the most critical, and main control board dispatches water quality sensor for data collection. LoRa RF unit is responsible for data interaction with the gateway.

The frame diagram of the main control board design is shown in Figure 2. Considering the complex water quality environment, multidimensional water quality data is collected for comprehensive analysis. For this reason, temperature sensor DS18B20, pH composite electrode E-201-C, turbidity sensor TSW-30, conductivity electrode DJS-1, and TDS sensor are selected to obtain the water temperature, pH value, turbidity, conductivity, and solids of the target water area. The main control chip adopts Atmega328p produced by ATMEL. This chip has rich peripheral interfaces such as timers, UART, ADC, I2C, GPIO, and SPI. And the built-in 64 kBytes Flash and 32 kBytes RAM can meet the access requirements of sensors and LoRa communication modules. In addition, the chip is based on ARM Cortex-M3 architecture, which has low power consumption and is suitable for long-term monitoring needs. The chip mainly completes data collection, processing, and storage and sends and receives data packets by the SX1278 RF unit. The data collection period of nodes is set to 30 min by the timer.

For this research conducted a more comprehensive analysis of water quality parameters such as water temperature, pH, turbidity, and conductivity. The selected sensor modules are shown in Figure 3. Among them, in water temperature-pH composite sensor, we use BNC interface and E-201-C type pH compos- ite electrode. In addition, the sensor has expanded DS18B20 temperature sensor interface. On the one hand, it can read the water temperature parameters, and on the other hand, it can compensate pH detection value to improve the accu- racy. The sensor uses 5 V working voltage and analog out- put. The working temperature is between 0-60°C, the measuring range is 0-14PH, and the response time is less than or equal to 1 minute.

The model of turbidity sensor selected in this study is TSW-30. The sensor comprehensively judges the turbidity by light transmittance and scattering rate in the target solu- tion. The sensor can output both analog and digital signals at the same time, and the working voltage is 5 V. The standard operating temperature is between -20°C and 90°C, and the detection response time is less than 500 ms.

Conductivity reflects the electrolyte concentration of the measured solution and is an important parameter to mea- sure the water quality. DJS-1 conductivity electrode in con- ductivity sensor is used for water quality monitoring. The sensor uses a 5 V supply voltage and a 0 ~ 3.4 V analog out- put. The working temperature is between 0 and 40°C, and the supported measurement range is 0-20 mS/cm.

The LoRa RF module is responsible for receiving and forwarding water quality data and works in 886 MHz unlicensed ISM frequency band in CAMBODIA. Main Control Unit (MCU) is programed to control the LoRa module and send information by the module after encoding. This design uses SX1278 baseband chip of Semtech Company. The chip uses LoRa chirp spread spectrum (CSS) technology to send data. This technology can eﬀectively reduce noise and interference and cover a large distance. Through an integrated +20 dBm power amplifier, it can ensure long-distance wire- less communication under the condition of sensitivity as low as -148 dBm. Moreover, the chip only costs about $1-2, which greatly reduces the design cost. Its super high-cost performance meets the management needs of equipment networking for remote water quality monitoring.

The circuit principle diagram of the LoRa communication module is shown in Figure 3. The communication module sx1278 is a RF module, which exchanges data with the ATmega328p, and RF antenna SMA connector packs the data and sends it to LoRa gateway. In order to ensure a better power supply performance, a magnetic bead is connected in parallel near the power supply of the sx1278 module. In the circuit design of RF antenna, use si9000 simulation calculation tool for PCB trace to control 50 Ω impedance of RF signal line. A π-type matching circuit is reserved to facilitate RF adjustment and ensure reliability.

**Gateway Design**

**Transport Layer Design**

**AI Prediction**

**Platform (Web/App)**

**Experiment and Results**

**Conclusion**

**Acknowledgments**

**References**