A Wireless Sensor Network Framework for Large-Scale Industrial Water Pollution Monitoring

Yohannes Derbew
yohanes.derbew@gmail.com

Mulugeta Libsie
Department of Computer Science, Addis Ababa University, Ethiopia
mulugeta.libsie@aau.edu.et

Abstract

Design and implementation of Wireless Sensor Networks (WSNs) is a challenging task because WSN research is usually application specific and each specific application requirement brings with it a different set of constraints and design objectives. Effective use of WSNs in large scale and long term monitoring of industrial water pollution requires satisfying two inherently contradicting requirements particular to this domain. On the one hand, it has to be time critical in order to provide early warning during compliance violations. On the other hand, it requires long-term data collection from a large number of distributed industrial sites for trend analysis to assist in decision making.

In this paper, we present the design of a comprehensive framework for using WSNs in industrial water pollution monitoring that aims to assist environment authorities in the decision making process of the battle against water pollution. The approach used to determine the suitability and effectiveness of our proposed framework is an experiment using a discrete-event simulator. The results show that our framework can provide in-network detection of compliance violations effectively without the need for powerful central processing stations thereby avoiding transmission of irrelevant data and consuming energy efficiently.

Keywords: Wireless Sensor Networks; Industrial water Pollution; Compliance Monitoring

1. Introduction

Water pollution is currently one of the major problems to the environment. With industrialization in major areas and urban cities growing, the water around them just keeps getting polluted. Recent research in Addis Ababa has classified Little and Great Akaki river basins as badly polluted to very badly polluted water [1].

There are a number of technologies developed to respond to the needs of effective monitoring to help prevent such pollution. Among these technologies, the most common option is the installation of industrial wastewater treatment plants that are capable of providing continuous and real-time monitoring services. However, these monitoring solutions require industrial users to install wastewater treatment plants that are very expensive and complex to operate.

This problem domain can benefit from WSNs. Compared with the present traditional monitoring techniques, enabling a monitoring system based on wireless sensor networks would present us with several advantages such as low cost, convenient monitoring arrangements, collection of a variety of parameters, high detection accuracy, high reliability of the monitoring results, etc.

Concerning that water is one of the pillars to sustainable economic development, a new, cost effective and efficient industrial water pollution monitoring solution using WSNs is proposed.

2. Background and Motivation

The Environmental Protection Authority of Ethiopia has standard procedures and policies to monitor and control the disposal of municipal and industrial wastes and by-products. Most of these policies and procedures are supported by manual practices that indirectly control the activities of plant and industry users through document reviews and field visits. These practices are not efficient because they do not directly monitor the quantity and range of contamination released from polluting industries. As a result, industrial users are becoming highly ignorant on the control of the wastewater they dispose into the environment.

It is insufficient to rely on the present traditional monitoring techniques to satisfy the current monitoring requirements, which emphasizes the fact that water pollution monitoring must be continuous, dynamic, macro-scale, and swift. Hence, to protect the environment from threats of factory water pollution, a higher degree of effectiveness is required in the monitoring and control of industrial wastewater discharges than those obtained through traditional monitoring.

Most current environmental monitoring solutions based on WSNs are characterized by generating a large amount of data if continuously used over a long period of time, which limits their application to our specific scenario where we have thousands of industrial sites each being monitored based on applicable compliance standards regarding wastewater discharges. In order to alleviate this limitation, there should be a mechanism to minimize the amount of data generated from each monitored site while maintaining the required compliance information.

On the contrary, environment authorities need to get and keep long term compliance information about each monitored industry. This requires sending and storing complete information from each monitoring network.

3. Related Work

A number of research activities have been reported for water quality monitoring in the literature. In [2], the authors proposed a monitoring system to measure parameters such as pH, dissolved oxygen, conductivity, and temperature for aquiculture, river, and lakes monitoring, where the convenience of using these kinds of systems in terms of price, flexibility, and real time processing is explained. The work simply involves only a single deployment site with instrumentation components tailored to address a relatively small number of monitoring parameters.

Another effort related to water monitoring WSN is exposed in [3] with a study of a river watershed in North Carolina. The authors presented an end-to-end hardware and software structure to monitor water parameters through a large number of highly distributed sensor networks.

The work in [4] is a study that shows how a system for water quality monitoring can be set up using Zigbee WSNs. The study proposes the use of high power Zigbee based WSNs for low cost, ad hoc and easy installation and maintenance. However, the system is proposed for a vendor specific platform, Zigbee, and the authors did not mention if it is possible to use the proposed system for a wide range of WSN platforms.

In [5], an expert system called WPC-ES is presented for assisting departments of environmental management in their efforts to improve water quality in a city, applied in the Yellow River Basin of China. The system can analyze relationships between industrial water pollution and economic activities of industrial enterprises of a city.

While the majority of these WSN solutions are focused in natural environments, the framework shown in this paper is based on the requirements imposed by the government for industrial pollutions in wastewater discharges. It would not be possible to apply the protocols and algorithms proposed in these solutions for monitoring a large number, possibly thousands, of distributed industrial sites.

4. System Requirement

Design of a WSN framework for effective use in large scale and long term monitoring of industrial water pollution requires meeting a set of requirements specific to the application scenario. We have identified the following set of basic requirements:

Energy Efficiency: Since batteries are currently the only power source for wireless sensor nodes, power conservation is an important issue that can be optimized based on the intended use of the sensor network. In our specific application scenario, the trade-off between delay and energy overhead, optimizations of energy consumption per reported event, and network lifetime have been considered.

Quality of Service: QoS attributes in WSN highly depend on the nature of the intended application. The design of our proposed framework aims to satisfy two important QoS attributes that need to be considered for this usage scenario. The first consideration is the goal of *near real-time detection* which requires our monitoring solution to detect an event of compliance violations and report as early as possible. We refer to the second consideration as *pollution event detection/reporting probability* which determines the effectiveness of our monitoring solution in detecting and reporting all possible violations all the time.

Compliance Monitoring: a solution based on WSNs for monitoring industrial wastewater discharges should be able to monitor a given industry based on applicable standard to make sure that only wastewater that complies with the standard is discharged into the environment.

Minimum Generated Data: When using our proposed framework for large scale monitoring, there would be thousands of distributed sensor networks installed on many industries all over the country each producing its own pollution related data. Unless we introduce some efficient strategies to reduce the amount of data generated and transmitted from each distributed sensor network to the central database, storing, analyzing, and managing the sensor data becomes difficult.

5. Solution Overview

To satisfy the mentioned design requirements, our proposed solution involves the design of a WSN framework having four main components:

i. Sensor Deployment Scheme

Our sensor deployment scheme is proposed with the following goals in mind:

- The deployment should be easily performed around the existing infrastructure with little or no need for physical modification across any kind of industrial facility.
- In order to make the monitoring framework generic and countrywide, sensor deployment should not be site specific.
- In order to detect illegal discharges effectively, the sensor nodes should be placed on points that are representative of the spatial extent of the wastewater.

ii. Network Architecture and Topology Design

Our proposed architecture offers a hierarchical tree topology that organizes the sensor network into a multi-hop network that can adaptively control the monitoring nodes under various pollution conditions that can happen in the wastewater effluent.

The topology is constructed by the exchange of routing layer advertisement beacons among all nodes in the network. A child node attaches itself to a parent node only if both have similar sensing modality as shown in Figure 1.

Our proposed topology construction protocol optimizes the famous Collection Tree Protocol (CTP), the “de facto of data collection in WSNs” [6]. We tailored this protocol to the needs of our specific application because:

- CTP does not allow a two way communication of application layer data.
- CTP does not use sensor modality as a metrics for parent selection.
- Routing layer data in CTP is not visible to the application layer for intermediate nodes to perform in-network data aggregation.

Figure 1 shows a typical network topology where sensors with similar modality connect to the same branch of the topology tree.

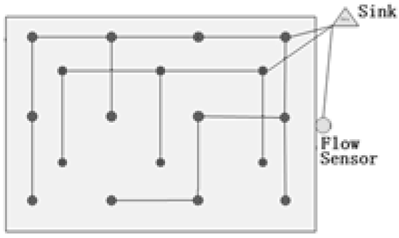

Figure 1: Sample network architecture and topology

iii. Context-aware Compliance Monitoring

Our proposed framework involves a rule-based node-centric context-aware model which reduces the amount of communication to the central server and improves energy efficiency. The following are the sources of context information that we consider:

- Conditions of pollutant parameters in the monitored wastewater: Pollutant concentration and Wastewater flow-rate
- Node role: (whether a node is a leaf or an intermediate node)

Our model is based on the Event-Control-Action (ECA) pattern presented in [7] for rule based context aware applications.

Figure 2 shows the high-level architecture of our proposed rule based node centric context-aware model.

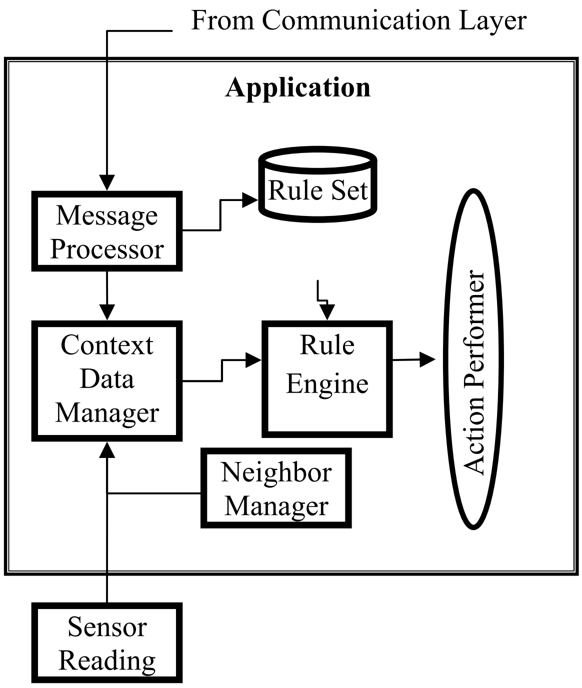

Figure 2: Architecture of the proposed node centric context aware compliance monitoring model

Message Processor identifies the type of message received from the communication layer and forwards the message to either the *Context Data Manager* which performs the task of capturing context information (about the wastewater and node role) from other nodes or locally from *Sensor Reading* as well as from the *Neighbor Manager* and provides temporary storage of this context information. The *Rule Engine* runs all the rules retrieved from the Rule Set to evaluate the context information coming from the local sensors as well as the context information gained from neighboring nodes. *Action Performer* is responsible for performing the actions triggered by the rule engine.

To cope with the limited local memory in the sensor nodes, the structure of the rules should be defined in order to express the compliance rules in a compact way as shown in Figure 3.

ID	Sensor Type	Condition	Min	Max	Action
(4)	(6)	(3)	(16)	(16)	(3)

Figure 3: Proposed rule format

The purpose of each field in the proposed rule format is discussed below:

ID (4 bits): a unique number that identifies each rule in the rule set.

Sensor Type (6 bits): is used to determine which sensor will be used in this rule. Each potential pollution parameter in the wastewater requires a specific type of monitoring sensor.

Condition (3 bits): the logical condition represents the current state of the wastewater such as current pollutant concentration, current role of a node, etc.

Min (16 bits), Max (16 bits): The values in these columns represent the limit values for each pollutant parameter as stated in the corresponding compliance standard.

Action (3 bits): Represent the operations of the rule that determine the reactions of the nodes in case a condition is true.

In order to reduce the computation at the sensor nodes, the contexts have to be described at the user side by defining a set of conditional rules which need to be applied for those contexts. After that, these rules will be translated into a form that the sensor node can understand using the proposed format.

For example, the ph of wastewater needs to be kept in the range of [5, 8] before discharging. Any value of the monitored wastewater ph out of this range has to be reported. For this scenario, the context rule can be set as follows:

IF ph OUT_OF [6, 8] THEN SEND READING

Here, the *ph* is represented by a value in the *Sensor Type* field; *OUT_OF* is represented by the *Condition* field; [6, 8] is represented by the *Min* and *Max* fields; and *SEND READING* is represented by the *ACTION* field.

iv. Environment aware Intra-Network Communication Protocol

In our proposed framework, the communication between the monitoring nodes and the sink node is environment aware that is based on the conditions of the wastewater effluent and builds upon the node-centric and rule-based context aware model that we proposed. It consists of four different phases: Initialization Phase, No-Pollution Phase, Pollution-Alert Phase, and No-Flow Phase.

A state diagram showing how the monitoring nodes and the sink node in the network transit between these phases based on environmental conditions in the wastewater effluent is given in Figures 4 and 5.

As can be seen from the state transition diagrams, each node switches from one phase to another based on the environmental conditions in the monitored wastewater. Each phase has a specific objective to achieve. The INITIALIZATION phase handles the exchange of control messages and context rules during network setup. The NO-POLLUTION and NO-FLOW Phases are designed to reduce network traffic and minimize energy consumption when pollutant concentration of the wastewater is in the acceptable range or when the flow rate of wastewater discharge is zero. The objective of the POLLUTION-ALERT phase is to ensure that our framework satisfies the near real-time detection goal.

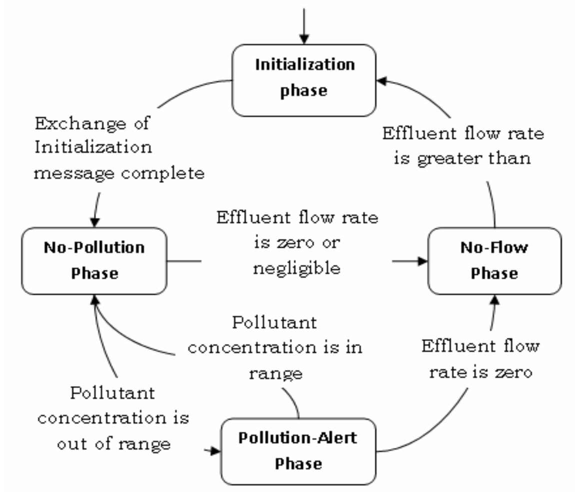

Figure 4: State Transition Diagram of a monitoring node

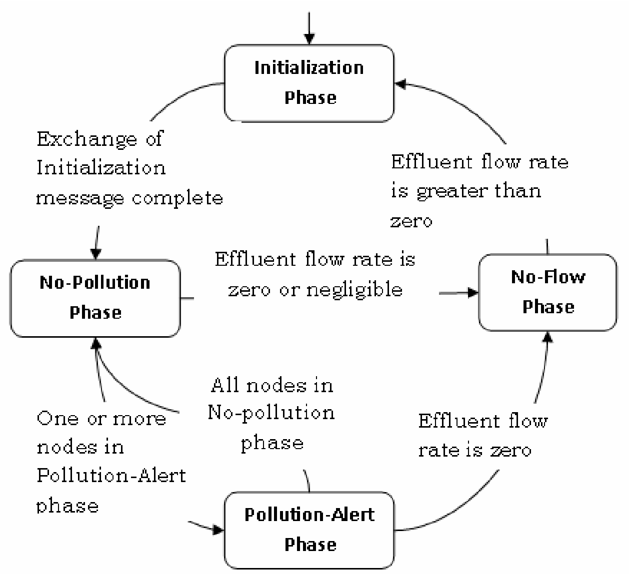

Figure 5: State Transition Diagram of the sink node

Figure 6 shows the format of the report message the monitoring nodes use to transmit non-complying pollutant concentration to the sink node during the POLLUTION-ALERT phase.

Node Count	Pollutant Type	Current Value

Figure 6: Format of the report message

The *Node Count* represents the number of nodes that have transmitted this report message so far. *Pollutant Type* is the index of the pollutant whose current value is being transmitted and *Current Value* represents the current average concentration for intermediate nodes and the current abnormal reading for the origin node.

Each intermediate node computes the average concentration of the pollutant using the aggregation function given below and sends out a single value to the next hop.

$$C_{avg} = \frac{\sum_{n=1}^{R}(T_n C_n) + C_{slf}}{N}$$, where

- C_{avg} = average pollutant concentration determined by the current node
- N = sum of the *Node Count* values of all report messages since last sample plus the current node
- R = total number of report messages received since last sample
- T_n = Node *Count value* of report message n
- C_n = value of the *Current Value* of report message n
- C_{slf}= concentration sensed by the current node

In the same manner, the sink calculates overall pollutant concentration in the wastewater periodically based on the report it receives from the monitoring nodes and sends a report message to interested parties only if the current concentration is out of the permitted range. The sink node uses the following aggregation function for this purpose:

$$C_{ocl} = \frac{1}{S}\sum_{n=1}^{N}(T_n C_n)$$, where

- C_{ovl} = overall pollutant concentration in the reservoir
- S = total number of sensor nodes in the network
- N = sum of all *Node Count* values in all received report messages
- T_n = *Node Count* in report message n
- C_n = *Current Value* in report message n

The third quantity the sink should determine during the POLLUTION-ALERT phase is the hourly *load* of the pollutant contained in the nonconforming wastewater effluent that is being discharged from the monitored factory. A pollutant *load* is the mass or weight of pollutant transported in a specified unit of time from pollutant source to the environment.

The following function is used to estimate the hourly pollutant load of the discharged wastewater effluent:

$$Load = K\sum_{n=1}^{N} C(\tau_n)Q(\tau_n)$$, where

- *Load* = hourly load of the pollutant under consideration
- K = a constant for converting units
- $C(\tau_n)$ = concentration of the pollutant during sampling period n
- $Q(\tau_n)$ = effluent discharge flow rate during the sampling period n
- N = number of samples taken in one hour

6. Experimental Results and Discussion

In order to prove the framework's applicability and efficiency, a model of the proposed framework is implemented using Castalia Simulator [8] that builds on the OMNeT++ simulation framework. The developed WSN model consists of 18 nodes that are organized into two groups. The first group contains 12 nodes (nodes 1 - 12) that monitor temperature. The remaining six nodes (nodes 13 -18) are used to monitor conductivity. Node 0 is the sink node in our model just as Castalia takes node 0 as a sink node by default. The deployment pattern is uniform among similar sensors and the deployment area is 50 x 60 meter.

We used Castalia's *Customizable Physical Process* model and made necessary changes to represent the variations in our monitored environment, which is *Temperature* or *Conductivity* of the wastewater. We also developed a separate application and routing layer messages formatted specifically to support our proposed network topology and communication protocols.

We performed simulations using different configurations to evaluate the various components of our framework and different design choices that we considered while deciding on a framework component. For each of the configurations, the simulation run is 100 minutes.

Scenario 1: Number of packets generated vs. node-centric context awareness compliance monitoring

In order to evaluate the effect of our context-aware model on the number of packets generated, we configured node 12 without context awareness (*Node12NoneContext*) while the rest of the nodes are configured with context aware operation (*ContextApplication*). The sink node is configured to collect data from each sensor node. The context rule used for this demonstration is a simple one. Each node transmits its sensor reading if the current reading is greater than a predefined report threshold. The number of packets received by the sink node from each sensor node, for each case is as shown in Figure 7.

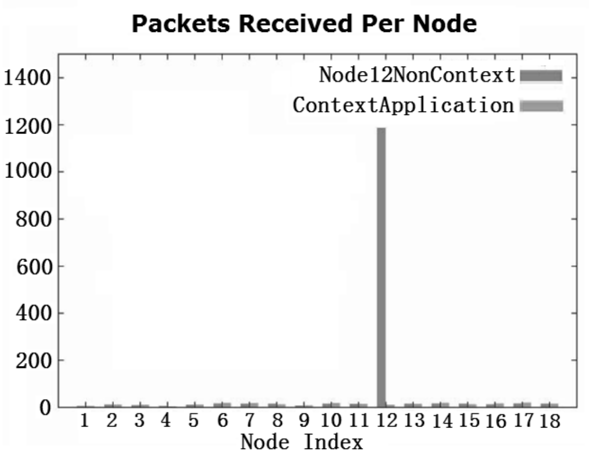

Figure 7: Number of packets received per node

Note that the sink received the largest number of packets from node 12 (which is run without our proposed context model) than from any other node during the simulation run.

In order to get better insight into the effect of using context aware sensor nodes for compliance monitoring, we performed a second simulation where all nodes, including node 12, are first run in context mode and then in non-context. The result of this simulation run is given in Figure 8. The graph shows the total number of packets generated by the whole network for the two cases.

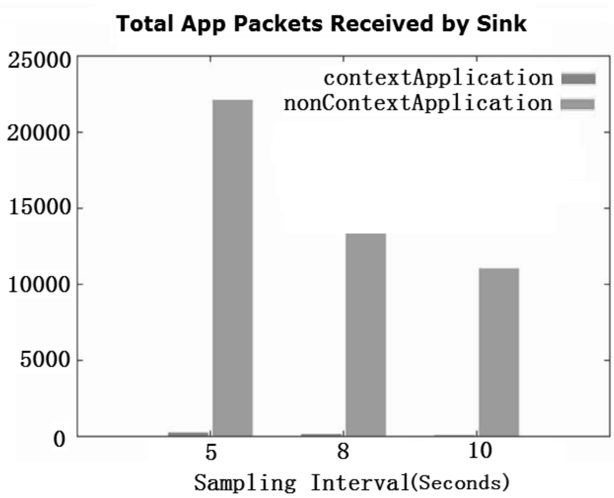

Figure 8: Effect of node centric context aware compliance monitoring on total network data traffic

Note from Figure 8 that it is also possible to reduce data traffic using larger sampling interval values. However, using longer sampling interval time increases the chance of missing events of hazardous pollutant discharge. This implies that our proposed compliance-monitoring model can reduce the amount of generated data even at higher sampling frequencies.

*Scenario 2: E*nergy consumption vs. node-centric context awareness compliance monitoring

As discussed before, by making sensor nodes aware of environmental context it is possible to reduce energy consumption of the sensor nodes and hence improve overall network lifetime. To show this, all nodes are first run in context mode and then in non-context mode using the same set of sampling rate values. The result of this simulation is presented in Figure 9, which shows the average energy consumption of the network in the two cases.

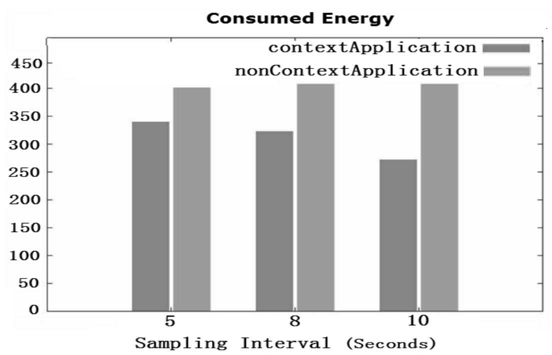

Figure 9: Effect of node-centric context aware compliance monitoring on energy consumption

Note from Figure 9 that the network consumes the maximum possible energy in the case of non-context operation even though we vary the sampling interval. However, in the case of context operation, the network can make considerable energy saving even for small sampling interval values.

Scenario 3: Number of packets collected vs. in-network data aggregation

Here, we evaluated our in-network data aggregation protocol based on its effect on the reduction of packets received by the sink node. To achieve this, we run all the sensor nodes alternately with the in-network aggregation protocol we developed and then without aggregation. The result of this simulation run is as shown in Figure 10.

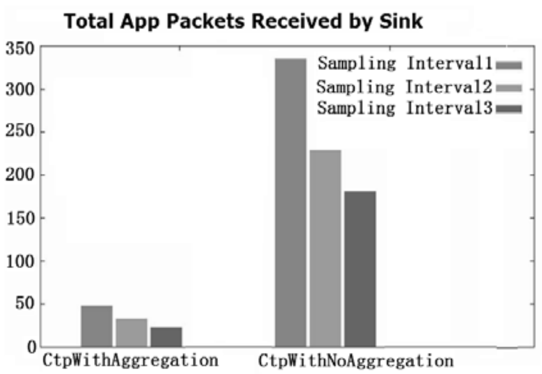

Figure 10: Effect of the proposed in-network aggregation protocol on network data traffic

The graph shows the number of packets received at the sink node during the simulation of 100 minutes both in the case of data aggregation and without it. Our proposed in-network aggregation protocol greatly reduces the amount of data the sink has to receive.

7. Conclusion

In this paper WSNs for use in large-scale industrial water pollution monitoring is presented. The outcome of the study is a novel WSN framework that incorporates four components each with a specific set of design objectives. The components of the proposed framework, along with their associated algorithms and protocols, are designed and optimized in order to satisfy the application specific requirements. The most important requirements that this work considered are the large number of distributed industrial facilities to be monitored, the possibility of future expansion and the need for industry specific compliance monitoring. Minimizing the amount of generated data traffic, energy efficiency, ability to monitor different industrial sites based on applicable compliance standards, and early detection are considered as design goals and treated in this work.

References

[1] G. Gebre and D. Van Rooijen, "Urban Water Pollution and Irrigated Vegetable Farming in Addis Ababa", 34th WEDC International Conference, Addis Ababa, Ethiopia, 2009.

[2] H.-B. Xia, P. Jiang, and K.-H. Wu, "Design of Water Environment Data Monitoring Node Based on Zigbee Technology", in Proceedings of CiSE '09, Wuhan, China, December 2009.

[3] G. W. Eidson, S. T. Esswein, and J. B. Gemmilletal, "The South Carolina Digital watershed: End-To-End Support for Real-Time Management of Water Resources", in Proceedings of the 4th International Symposium on Innovations and Real-time Applications of Distributed Sensor Networks (IRADSN '09), Vol. 2010, May 2009.

[4] Z. Rasin and M. Rizal Abdullah, "Water Quality Monitoring System Using Zigbee", International Journal of Engineering and Technology IJET Vol. 9, No. 10, 2010.

[5] Cheng H., Yang Z., and Chan C. W, "An Expert System For Decision Support of Municipal Water Pollution Control", Engineering Applications of Artificial Intelligence, Vol. 16, No. 2, 2003.

[6] O. Gnawali and R. Fonseca, "Collection Tree Protocol", ACM, Berkley CA, 2009.

[7] L. Maria Daniele, "Towards a Rule-based Approach for Context-Aware Applications", Unpublished Master's Thesis, University of Cagliari, Italy, 2006.

[8] Athanassios Boulis, "Castalia: a Simulator for Wireless Sensor Networks and Body Area Networks", User's Manual, Version 3.2, 2011.