7-20-2016

System-Aware Smart Network Management for Nano-Enriched Water Quality Monitoring

Bassem Mokhtar
Alexandria University

Mohamed Azab
Virginia Polytechnic Institute and State University

Nader Shehata
Utah State University

Mohamed Rizk
Alexandria University

Follow this and additional works at: https://digitalcommons.usu.edu/biology_facpub

Recommended Citation
Mokhtar, Bassem; Azab, Mohamed; Shehata, Nader; and Rizk, Mohamed, "System-Aware Smart Network Management for Nano-Enriched Water Quality Monitoring" (2016). Biology Faculty Publications. Paper 1173.
https://digitalcommons.usu.edu/biology_facpub/1173

This Article is brought to you for free and open access by the Biology at DigitalCommons@USU. It has been accepted for inclusion in Biology Faculty Publications by an authorized administrator of DigitalCommons@USU. For more information, please contact dylan.burns@usu.edu.
Research Article

System-Aware Smart Network Management for Nano-Enriched Water Quality Monitoring

B. Mokhtar, 1,2 M. Azab, 2,3,4 N. Shehata, 2,5,6 and M. Rizk 1,2

1Department of Electrical Engineering, Faculty of Engineering, Alexandria University, Alexandria, Egypt
2Center of Smart Nanotechnology and Photonics (CSNP), SmartCI Research Center, Alexandria University, Alexandria, Egypt
3Informatics Research Institute, City of Scientific Research and Technological Applications, Alexandria, Egypt
4ACIS, ECE Department, University of Florida, Gainesville, FL, USA
5Department of Engineering Mathematics and Physics, Faculty of Engineering, Alexandria University, Alexandria, Egypt
6USTAR Bioinnovations Center, Utah State University, Logan, UT, USA

Correspondence should be addressed to B. Mokhtar; bassem.mokhtar@gmail.com

Received 18 April 2016; Revised 19 July 2016; Accepted 20 July 2016

The detection of water quality parameters such as dissolved oxygen (DO) and potential hydrogen (PH) in aqueous media is important for a wide variety of applications including environmental monitoring, biomedical research, and process control [1–3]. Compared to currently used techniques, fluorescence-based sensing technique has significant advantages over other procedures due to fouling avoidance, the fact that there is no need for a reference electrode, and the resistance to external electromagnetic field interferences [4–9]. One of the most promising optical nanostructures is ceria nanoparticles due to its oxygen capability storage, low-cost synthesis, and adequate sensitivity [10, 11]. This paper offers a comprehensive monitoring framework as an integration between nanotechnology and trustworthy wireless sensor networks.

1. Introduction

The main objective of our work is to develop a complete sensing platform for real-time monitoring of two water quality parameters, DO concentration and PH value, in aqueous media [12, 13]. Our system, as shown in Figure 1, goes behind local, single location monitoring to a networked sensing of DO and PH concentrations at multiple locations, across streams, water treatment facilities, hydroponic farms, and aquafarms. The system comprises two main subsystems, which are the data sensing and forwarding subsystem (DSFS), and Operation Management Subsystem (OMS). The DSFS employs a Hybrid Intelligence (HI) scheme which is proposed through integrating an association rule learning algorithm with fuzzy logic and weighted decision trees. The DSFS operation is based on profiling and registering raw data readings, generated from a set of optical nanosensors, as profiles of attribute-value pairs. As a case study, we evaluate our implemented test bed via simulation scenarios in a water quality monitoring framework. The monitoring processes are simulated based on measuring the percentage of dissolved oxygen and potential hydrogen (PH) in fresh water. Simulation results show the efficiency of the proposed HI-based methodology at learning different water quality classes.
forwarding processes. Powerful nodes receive the digitalized signal from one or more running nanosensors, merge the collected data with the geographical location of the sensing element(s), and run the HI techniques for learning data patterns and taking on-site decisions. Based on the weight of reached decisions, data might be forwarded to sinks and remote off-site management servers in a real-time fashion for further analysis with security extensions. The real-time management of the overall system with security policies’ enforcement is performed by the OMS. The OMS operates based on learned patterns and rules of system operations projected from the DSFS to manage the entire network of sensors. The main tasks of OMS are to enable real-time data visualization, managed system control, and secure system operation. As shown in Section 3, the main advantage of having such level of autonomic control and management is the ability to predict some of the sensor feedback based on the analysis of the overall network feedback [14–16]. Such ability facilitates optimizing the sensor power usage by controlling the activation periods of sensors leading to much better sensor utilization expanding the lifetime of the entire network and reducing the system cost. Further, such ability facilitates detecting and fixing/excluding any misbehaving or problematic sensing nodes that can massively reduce the system accuracy. Additionally, the automated system could definitely overcome the problems of manual data collection that requires extensive effort and time. This capability enables almost continuous real-time monitoring of DO with means to identify and address sensor drift helping researchers to construct test models based on the flow of the monitored water supply. Both DO and PH are sensed using fluorescence quenching technique depending on the reduction of visible emitted fluorescent optical intensity based on the change of the quencher concentration of dissolved oxygen or the chemical impact of PH in ceria nanoparticles itself. Additionally, we present in this paper a HI-based scheme for data classification and forwarding in wireless sensor networks (WSNs) enabling energy-efficient, better utilized resources and optimized QoS operations. The main objectives of our proposed scheme, by adopting concepts in the information theory and Artificial Intelligence (AI) [17–19], are as follows:

(i) Building efficient reconfigurable information-driven data classification and forwarding methodology for WSNs.

(ii) Learning interesting routing-based attributes and generating forwarding models.

(iii) Enabling the capability of learning/expecting/detecting abnormal data flows and making classification and generating rules.

The organizing of the remaining sections of this paper is as follows. Section 2, Materials and Methods, discusses three subsections including the methodology of data sensing, forwarding subsystem with the proposed HI-based data classification/forwarding scheme, the experimental setup for sensing of DO and PH optical nanosensors, and the trustworthy Operation Management Subsystem. The obtained results are analyzed and discussed in Section 3. Section 4 concludes the paper and highlights our future work.
2. Materials and Methods

2.1. Data Sensing and Forwarding Subsystem (DSFS). The networking infrastructure of sensors’ communication and data forwarding processes within the DSFS depends on having a network of wirelessly communicated sensors organized in clusters. The continuously growing interest and developed research in Wireless Sensor Networks (WSNs) lead to making such networks widely used in applications related to various fields, such as environmental monitoring [20]. WSNs provide the communication framework for such applications enabling data sensing, collection, and forwarding until reaching the final data destination center for further data analysis and decision-making. As known, WSNs comprise limited energy and computation communicating nodes; thus, this leads to great challenges on the lifetime of such networks and consequently their reliability. So, there is a need to propose an efficient data forwarding scheme that considers the energy constraints in WSN and the importance of interesting data in order to route data that have large information gain. In the literature, many trials provided solutions for having efficient data forwarding and routing targeting energy saving at sensor nodes using opportunistic routing theory [21, 22]. Other researchers consider data routing and forwarding depending on data aggregation and clustering approaches [23]. Another work provides an effective data collection using a set of powerful mobile nodes which increases the probability of having available and reliable communication channels [24].

The inter distance and noise level between sensor nodes is sometimes considered for designing data forwarding strategies [25]. However, to the best of our knowledge, prior related work did not provide a data classification and forwarding scheme for WSNs leveraging the capabilities of simple hybrid intelligence techniques and information theory concepts to extract data semantics and learn interesting information. Accordingly, there will be an ability to control classification levels and amounts of data forwarded from areas of interest to remote data analysis and decision making locations. This will, consequently, aid in optimizing power consumption and memory utilization at sensor nodes. Leveraging concepts form information theory; this paper offers a novel hybrid intelligence-based scheme for data classification and forwarding in wireless sensing communication environments enabling energy-efficient, better utilized resources and optimized QoS operations.

As a case study, we run this scheme for a WSN directed to monitor water quality levels. The proposed methodology depends on profiling raw data readings, generated from a set of optical nanosensors, as profiles of attribute-value pairs. Then, data patterns are learnt adopting association rule learning algorithm clarifying the most frequent attributes and their related values. According to the discovered sets of attributes, a set of fuzzy membership functions are directed to produce a discrete sample space and a specific membership class for each attribute based on its value. Based on calculated attribute-dependent entropies and information gains, weighted probabilistic decision trees are built to help take decisions of data forwarding and to generate long-term fuzzy rules. One of the main objectives of our proposed scheme, by adopting concepts in the information theory and Artificial Intelligence (AI), is building efficient reconfigurable information-driven data classification and forwarding methodology for WSNs [17–19].

2.1.1. Hybrid Intelligence-Based WSN-Based Data Classification and Forwarding Scheme. Figure 2 shows the building blocks of the proposed scheme for data classification and forwarding for WSNs. Those blocks illustrate the main used algorithms and operations. We will discuss the comprised blocks shown in Figure 2 as follows.

(i) The shared memory: an accessible memory located at sensor nodes with powerful storage and processing resources. Data readings are stored as profiles of attribute-value pairs. Each data profile refers to a specific reading from a certain sensor node located at a location where that sensor is directed to execute a defined function, such as reading lead-related quality level at fresh water.

(ii) Association rule learning algorithm: a machine learning algorithm, which is adopted to learn data pattern at the shared memory and to extract the most frequent attributes located in the memory according to defined minimum support and confidence thresholds. Those attributes will be forwarded to the fuzzifiers in order to be classified based on attributes’ values.

(iii) Fuzzifiers: operating engines which adopt fuzzy logic via defining a set of defined fuzzy membership functions to generate specific finite discrete classes for the attributes based on their values [26].

(iv) Statistical analysis: this defines a set of statistical operations that are performed on data profiles stored in the shared memory in order to aid in extracting some information leading to computing the information gains. For instance, a specific attribute will be analyzed to know how many times it is found in readings related to pollutants based on measured levels of dissolved oxygen and lead in fresh water at using specific types of sensors.

(v) Entropy and information gain calculation: it runs mathematical operations based on information concepts to calculate the information entropy due to classified attributes and the possible outcome classes after decision making. Then, it computes the information gains for each attribute with respect to the calculated class entropy. Accordingly, the attribute which has the greatest weight on making decisions will be known. Then, other attributes with weights ordered in a descending order will be used to form the decision tree. The next section will discuss an illustrative example to show how calculations are performed.

(vi) Decision trees and fuzzy rules generation: based on the above calculations, weighted entropy-based decision trees will be formed and used to make decision. According to these designed decision trees and possible attribute classes, a set of fuzzy rules can
be generated which shows the different cases that might be faced when getting a set of attributes within certain range of values.

2.1.2. Overview of the Technical Insights of the Proposed Scheme for Efficient Water Quality Monitoring. In this subsection, we will discuss the main concepts on which the proposed hybrid intelligence scheme for data classification and forwarding is developed. We assume a small-scale WSN comprising a set of communicating sensor nodes where some of those nodes, called advanced nodes, are with powerful computing and storage capabilities. This WSN is directed to collect data related to various classes of water pollutants and forward data to a final destination unit in order to make further data analysis and management. Collected data and readings from sensor nodes are stored in a shared on-demand accessible memory, where data entries are indexed according to sensors’ identifications and their locations. Advanced nodes can play the role of cluster heads in case of forming a cluster hierarchy for the designed WSN. Also, advanced nodes analyze the patterns of stored data, where these data refer to readings about a certain pollutant. The following subsection will discuss building weighted classifiers, using fuzzy logic, entropy, and decision trees, which are adopted by sensor nodes [27].

The following points illustrate how the weighted classifiers are built and work in advanced sensor nodes. As discussed previously, readings from sensors are stored as attribute-value pairs in a shared memory.

(i) The fuzzy logic divides the registered attributes with an infinite real space into finite sets with discrete finite space.

(ii) According to obtained attributes’ levels and expected related classes, a statistical analysis can be done which will show the percentage/probability of each attribute to a possible class.

(iii) Based on data pattern and using the entropy of all possible classes and the conditional entropy with respect to their related attributes, a fuzzy weighted decision trees can be built. Equation (1) calculates the class entropy:

$$H(C) = -\sum_{C=1}^{n} p_C \log_2 p_C,$$

where $p_C$ shows the probability of a certain class found in the analyzed patterns. It can be calculated based on the frequency of a certain class in data patterns. For example, if there are 10 data profiles where 5 of them refer to a water quality class of high level type, then the probability of high level quality class is 50%. Then, to know the weight of each attribute that will affect getting the main class, we will use the conditional entropy as described in

$$H(C/A) = \sum_{a \in A} (p_a) H(C/A = a),$$

where $a$ is a possible value for attribute $A$ and $H(C/(A = a))$ is the conditional entropy of having a certain class when attribute $A$ is of value $a$.

Assume that we have three fuzzy attributes (i.e., $A = 3$) that can affect the process of classification. We need to know the more affecting attribute that will be the first stage the classifier. This can be obtained from calculating the conditional entropy for all attributes and their possible values which can be defined from the fuzzifiers by which the attributes are with a finite sample space. For instance, there is an attribute $A_{type}$, such as sensor type, which can take two
different string values (i.e., $a = 2$), which are $X$ and $Y$. This is a binary attribute. From patterns, we found that the value of 20% of patterns which contain $A_{type}$ is $X$. Then, $(p_{a=X}) = 0.2$ and $(p_{a=Y}) = 0.8$.

Then, we will use (1) and (2) to calculate the information gain as stated in

\[
\text{Information Gain} = I = I(C;A) = H(C) - H\left(\frac{C}{A}\right). \quad (3)
\]

Based on the calculated gain, we can know the attributes with higher priorities and those attributes will be at the first stages or nodes in the decision tree-based classifier. The more the information gain we will have for an attribute, the higher the probability that such attribute will be in the first decision tree nodes.

For example, if $H(C) = 1.5$ bits and we have three attributes, which are $A_{type}$, $A_{level}$, and $A_{status}$, and each attribute has two possible string values, where $H(C/A_{type}) = 1.1$ bits, $H(C/A_{level}) = 0.7$ bits, and $H(C/A_{status}) = 0.5$ bits, this means that information gain according to $A_{type}$ is the largest one. So, the first node in the decision tree will be the sensor type attribute. Figure 3 shows an example of a decision tree:

(i) Some fuzzy rules from the built decision trees can be extracted [28]. Relying on those rules, data forwarding and routing decisions can be taken. For instance, a rule of a low level of quality can be extracted based on the designed decision tree as shown in Figure 3. As an example, if we have a set of ON dissolved $O_2$ sensors with high level readings, then this refers to a low quality level.

As a case study with an illustrative example, we discuss a numerical example which discusses the designed forwarding scheme. We assume that WSN is employed to forward quality-related data readings from a set of sensors to a final destination in a fresh water context like a river. The implemented WSN will comprise two main types of sensors, which are normal sensors and advanced sensors. The proposed scheme will be implemented in advanced sensor nodes, which possess powerful computing resources and communication capabilities. Advanced nodes will be able to register readings from normal nodes in a shared memory at them allowing other nodes to access and get more information (e.g., learning data patterns concerning a certain pollutant during a specific period in the year). Readings will be represented by advanced nodes in their memories as data profiles of attribute-value pairs. The data profile will show the identification (ID) of the data source sensor node, type of sensor node, status of the sensor (ON, OFF, and maintenance), and level of readings (e.g., high or low).

Normal nodes will be equipped with two types of sensors that can provide readings about the percentage of dissolved oxygen and $PH$ as indication to the quality of the fresh water in the studied river. The practical measurements of dissolved oxygen in fresh water show the following concentrations
Concerning the concentration level of PH in a fresh water, we adopt the following three levels as indication to having acidic levels [29]:

(a) Low acidic (6–8).
(b) Moderate acidic (4–6).
(c) Highly acidic (<4).

The designed scheme in advanced nodes will adopt fuzzy logic that employs fuzzy membership functions (FMFs) to classify the readings levels of each sensor. For readings concerning the dissolved oxygen and the discussed concentration levels previously, FMFs are designed, as shown in Figure 4, to consider the following: the first concentration as best quality, the second concentration as moderate quality, and the last two concentrations as low quality. For the concentration level of absorbed lead, other FMFs are used to define the first concentration level as safe level or optimum quality of water and the second concentration level as moderate quality and the last concentration level as the lowest quality of water.

In this example, we assume that we have one final destination, four advanced nodes, and five normal nodes. Also, we have two types of sensors in each sensor node for measuring levels of DO concentration and PH. The data profile of a reading stored in a memory of an advanced sensor node, type of sensor, reading level, and status of the node.

Considering that each region in a river is covered by a set of sensor nodes with specific IDs for monitoring water quality, so each area will have data patterns (based on registered data profiles) that can be learned to know and detect the main pollutants that might exist and the expected consequences and recommend countermeasure for treatment.

![Figure 4: FMFs for dissolved oxygen concentration.](image)

**Table 1: Data profiles of sensors.**

<table>
<thead>
<tr>
<th>ID</th>
<th>Sensor type</th>
<th>Reading level</th>
<th>Sensor status</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DO</td>
<td>High (&lt;3 ppm)</td>
<td>ON</td>
<td>Dangerous</td>
</tr>
<tr>
<td>1</td>
<td>PH</td>
<td>Low (5)</td>
<td>ON</td>
<td>Moderate</td>
</tr>
<tr>
<td>2</td>
<td>DO</td>
<td>No (6–8 ppm)</td>
<td>ON</td>
<td>Optimum</td>
</tr>
<tr>
<td>2</td>
<td>PH</td>
<td>Low (4)</td>
<td>ON</td>
<td>Moderate</td>
</tr>
<tr>
<td>3</td>
<td>DO</td>
<td>High (&lt;3 ppm)</td>
<td>ON</td>
<td>Dangerous</td>
</tr>
<tr>
<td>3</td>
<td>PH</td>
<td>N/A</td>
<td>OFF</td>
<td>Maintenance</td>
</tr>
<tr>
<td>4</td>
<td>DO</td>
<td>High (&lt;3 ppm)</td>
<td>ON</td>
<td>Dangerous</td>
</tr>
<tr>
<td>4</td>
<td>PH</td>
<td>Low (6)</td>
<td>ON</td>
<td>Moderate</td>
</tr>
<tr>
<td>5</td>
<td>DO</td>
<td>High (4 ppm)</td>
<td>ON</td>
<td>Dangerous</td>
</tr>
<tr>
<td>5</td>
<td>PH</td>
<td>No (8)</td>
<td>ON</td>
<td>Optimum</td>
</tr>
</tbody>
</table>

In normal operation, we get two readings from each normal sensor node through the day. So, we expect 10 readings per day and 300 readings per month. As shown in Table 1, assume that we have for a region of interest in the river the following daily data profiles collected from five operating sensors through a certain month.

The advanced sensor nodes will learn the patterns of data profiles stored in their memories. Advanced nodes will learn Apriori-based association rule learning algorithm [30] for learning the main frequent attributes in the stored data profiles, which are sensor ID, sensor type, reading level, and sensor status. Additionally, advanced nodes will employ statistical analysis algorithm for generating some statistics about stored data profiles, such as how many profiles contain O₂ and high pollution level [31].

From Table 1, advanced nodes can build a weighted decision tree-based on entropy and information gain in order to classify data flows and know the more important flows. Hence, they can allocate more resources and offer more bandwidth.

For building the weighted decision tree, we will have the following procedures. There are 10 data profiles with three main attributes, which are sensor type, reading level, and sensor status. Also, there are main four decisions in the table which are maintenance, no pollution, low pollution, and high pollution. From (1), we can calculate the decision entropy as follows:

\[
H(D) = - \sum_{D=1}^{5} p_D \log_2 p_D
\]

\[
= -p_{\text{Dangerous}} \log_2 p_{\text{Dangerous}} - p_{\text{Moderate}} \log_2 p_{\text{Moderate}} - p_{\text{Optimum}} \log_2 p_{\text{Optimum}} - p_{\text{Maint}} \log_2 p_{\text{Maint}}
\]

\[
= -0.4 \log_2 0.4 - 0.3 \log_2 0.3 - 0.2 \log_2 0.2 - 0.1 \log_2 0.1
\]

\[
= 0.529 + 0.521 + 0.4644 + 0.3322 = 1.8466 \text{ bits.}
\]
Then, we will use (2) and (3) to compute the information gain for each interesting attribute of the main three we have.

For the first attribute: sensor type

\[
I(D; A) = H(D) - H\left(\frac{D}{\text{Type}}\right)
\]

\[= 1.8466 - \frac{|O_2|}{10} \text{Entropy}(O_2) - \frac{|PH|}{10} \text{Entropy}(PH)
\]

\[= 1.8466 - 0.5 \times (-0.8 \log_2 0.8 - 0.2 \log_2 0.2)
\]

\[= 1.8466 - 0.5(0.2575 + 2 \times 0.4644)
\]

\[= 1.8466 - 0.8 = 0.8 \text{ bits.}
\]

(5)

Similarly for the other two attributes,

\[
I(D; A) = H(D) - H\left(\frac{D}{\text{Level}}\right) = 1.8466
\]

\[= 1.8466 - \frac{|High|}{10} \text{Entropy}(High) - \frac{|Low|}{10} \text{Entropy}(Low)
\]

\[= 1.8466 - 0.4(0) - 0.3(0) - 0.2(0) - 0.1(0)
\]

\[= 1.8466 \text{ bits.}
\]

(6)

\[
I(D; A) = H(D) - H\left(\frac{D}{\text{Status}}\right) = 1.8466
\]

\[= 1.8466 - \frac{|ON|}{10} \text{Entropy}(ON) - \frac{|OFF|}{10} \text{Entropy(OFF)}
\]

\[= 1.8466 - 0.9(-0.444 \log_2 0.444 - 0.333 \log_2 0.333)
\]

\[= 0.222 \log_2 0.222 - 0.1(0) = 1.8466 - 0.9(0.52)
\]

\[= 0.5283 + 0.482 = 1.8466 - 0.9(0.566256)
\]

\[= 1.337 \text{ bits.}
\]

From the above results, we find that the higher information gain is obtained for the reading level attribute (1.8466 bits) and then for the sensor status (1.337 bits) and then for the sensor type (0.8 bits). This means that the advanced powerful sensor nodes will forward data to the final destination management units once they detect high reading level from operating sensor nodes without giving more weight to the sensor type or the main cause for the low water quality levels. Accordingly, a weighted decision tree can be formed and adopted by advanced nodes to take decision and forward data to the final destination. A set of rules can be generated according to the built decision tree such as the following: if there is a high level reading from an operating DO or PH sensor, then forward the data to the destination.

2.2. Optical Materials’ Synthesis and Sensing Setup. Ceria nanoparticles are prepared using a chemical precipitation technique for simplicity and relative low-cost precursors [32–34]. 0.5 g of cerium (III) chloride (heptahydrate, 99.9%, Aldrich Chemicals) is added to 40 mL deionized water as a solvent. The solution is stirred at rate of 500 rpm for 2 hours at 50°C with added 1.6 mL of ammonia. Then, the solution is stirred for 22 hours at room temperature. The synthesized ceria nanoparticles are characterized using UV-Vis spectroscopy (dual beam PG 90+) to detect the absorbance dispersion and consequent bandgap calculation.

The experimental test bed used to detect the change of the fluorescence intensity peak due to the change of DO concentration and PH value in the aqueous solution is shown in Figures 5 and 6. The fluorescence spectroscopy system consists of a near-UV LED of central wavelength of 430 nm as an excitation source exposed to a three-neck flask containing DI water solution of the synthesized ceria nanoparticles. Oxygen and nitrogen gases are fed through individual lines through double-holes cork placed into one of the necks on
the flask and controlled by a mass flow rate controller. The probe of a commercial DO meter (Thermo Scientific A500 with a measurement range up to 50 mg/L) is inserted in the second neck of the flask to measure DO concentration. The fluorescence signal is collected from the colloidal solution scanned by a Newport Cornerstone 1300 monochromator, positioned at a 90° angle to the excitation signal for minimum scattering effects. Then, a photomultiplier tube (Newport PMT 77340) is connected to a power meter (Newport Power meter 1915-R) for fluorescence intensity monitoring. Same setup is used for PH detection at normal DO concentration, by removing both oxygen and nitrogen inlets with varying added concentrations of acid (HCL). The value of PH is measured using same meter (Thermo Scientific A500) but with the optimum probe for PH detection.

2.3. Autonomous Sensor Operation Management System. Figure 7 is a block diagram of the sensor framework where the main sensing element is interfaced with an autonomously managed wireless sensor network for water quality monitoring. The sensing function is interconnected to the cyber layer for control, management, and monitoring.

The nanosensor is interfaced with an A2D chip, a powerful microcontroller chip, a GPS chip, and wireless transmission module. The microcontroller is programed to control the activation and deactivation of the sensing element and control the measurement configuration if necessary. The microcontroller receives its guideline from a remote management server. Each sensor has a unique identifier that is used in all transmissions along with the geographical location of the sensor in case of mobility. The system is built to be as generic as possible allowing more sensing elements to be attached to the same sensing elements if the application requires that.

The system is built to scale, where sensors are grouped into different zones. The sensor feedback and location are dispatched frequently based upon either event change and query or a predetermined schedule to a central data collection node at each zone. This node applies partial analysis and data grouping and dispatches a comprehensive zone status-report to the management server for further analysis and guidance. The algorithms used to establish such analysis are application dependent. We devised a simple case study with a simple model for the excremental study just to reflect the effect of such automation on the quality of the system output.

3. Results and Discussion

3.1. Evaluation of HI-Based Data Classification and Forwarding Scheme. In this subsection, we conduct a simulation.
3.1. Simulation Setup. In this section, we provide a preliminary simulation study for the proposed classification and forwarding scheme. Using the discussed example in Section 2.1.2, we conducted a simple scenario of a WSN-based pollution monitoring and data forwarding network using Java based discrete time event-driven WSN simulator, called J-Sim [35]. We target a river as a network context where sensors will be distributed along a river bank. Figure 8 shows a layout of the scenario which depicts the network topology and communication amongst sensor nodes. We assume a linear WSN where sensors forward data to the next available sensors. We have two different types of sensors which are the normal node and the advanced node. The last type can access registered data patterns and decide whether the forwarded data are of a certain importance or not. According to the calculated information entropy and gain besides the constructed weighted binary decision trees, the advanced nodes will forward data to the final network destination. We assume that we have two types of pollutants that will result in various levels of pollution indicators which depend on measuring the concentrations of dissolved oxygen (DO) and PH levels in regions of interest in the river. In case of having low measurements of DO and low levels of PH, this means that we have indication of a combined low water quality level. On the other hand, if there are measurements’ levels, from one sensor type, referring to a certain low water quality indication, this means that we have an explicit low water quality level. Table 2 shows the simulation parameters. Table 3 shows a set of key metrics which are used to study the offered QoS and network performance.

Figure 9 shows that advanced nodes allocate more resources and bandwidth for data flows concerning high level readings from dissolved O$_2$ sensors. The figure shows that, at early time, low water quality levels are detected and then huge amount of data are allowed to flow from normal sensor nodes to the destination. Also, some advanced nodes find high water quality levels based on measured dissolved O$_2$ concentrations. Once the water quality-related data patterns reach a certain information gain, advanced nodes allocate more resources and forward the related data to the final destination. So, in case of optimum water quality levels, less traffic is found since small information gain is calculated (i.e., we have in this case normal operations with low information entropy).

Figure 10 shows that there are higher data throughputs measured at the final destination in case of detecting low water quality levels in case of having one source (only dissolved O$_2$ concentrations) or with two sources (dissolved O$_2$ concentrations and PH levels). This is because more data packets will be forwarded amongst sensor nodes till reaching the destination.

Figure 11 depicts the response of the WSN to a low water quality level that exists widely in the studied context.
Many advanced nodes are able to detect low dissolved O$_2$ concentrations at the studied context and, hence, they allow data flows to pass to the final destination.

We studied the average residual power at the first dead sensor node. As shown in Figure 12, the degradation in power increases as there are readings related to low water quality levels. Hence, more data packets will be forwarded to the final destination. Also, the figure shows that combined low water quality levels, which comprise readings related to low quality indicators based on the measured concentrations of dissolved O$_2$ and PH levels, result in more degradation in...
Table 3: QoS and network performance metrics.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Unit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Received data throughput</td>
<td>Bytes/seconds</td>
<td>The amount of data received by the final destination</td>
</tr>
<tr>
<td>Average maximum throughput</td>
<td>Bytes/seconds</td>
<td>The average of maximum throughput captured at the final destination when running the same scenario for more than one time</td>
</tr>
<tr>
<td>Water quality level detection time</td>
<td>Seconds</td>
<td>Delay introduced by the time needed to measure amounts of data readings for determining a specific water quality level</td>
</tr>
<tr>
<td>Average residual power per sensor</td>
<td>Watt</td>
<td>The average amount of measured residual power at sensors after running the simulation scenario and forwarding data</td>
</tr>
</tbody>
</table>

3.2. Optical Nanoparticles Characterization. Figure 13(a) shows the resulting absorbance dispersion of the synthesized nanoparticles. Based on the resulting absorbance measurements, the allowed direct bandgap semiconductor of the synthesized nanoparticles can be found through the following equation [36]:

$$\alpha E = A^* (E - E_g)^{1/2},$$  \hspace{1cm} (7)

where $A^*$ is a constant for the given material depending on its refractive index and effective masses of both electrons and holes, $E$ is the absorbed optical energy, and $E_g$ is the direct allowed bandgap energy. Then, $(\alpha E)^2$ is plotted with the absorbed optical energy, and the intersection with $x$-axis gives the value of bandgap energy as shown in Figure 13(b).

Figure 14 shows the change of the visible fluorescence emission intensity at 520 nm from the ceria nanoparticles with increasing DO concentration at normal PH, ~7, under

Decrease DO in ceria in normal water

![Decrease DO in ceria in normal water](image)

Figure 14: Visible fluorescence spectra at different DO concentration within neutral PH (~7).
near-UV excitation. We speculate that this is due to a release of oxygen stored in the ceria lattice when the nanoparticles are introduced into the solution. Also, the emitted fluorescence emission is found to be reduced with increasing the added acidic concentration which consequently reduces the value of PH as shown in Figure 15, at normal DO concentration level (~8 ppm).

4. Conclusions

This work presented a smart network management system for efficient data classification and forwarding and decision making. The system comprised two main subsystems, a data sensing and forwarding subsystem (DSFS) and Operation Management Subsystem (OMS). The DSFS adopted a novel hybrid intelligence (HI) scheme for data classification and forwarding in wireless sensor networks incorporating information theory concepts, machine learning, fuzzy logic, and weighted decision tress. Such adoption led to better energy consumption, improved resource utilization, and optimized QoS operation. Simulation scenarios discussed the performance of a WSN employing the proposed scheme for monitoring water quality indicators. Simulation results of the proposed HI-based data classification and forwarding scheme showed that the DSFS works efficiently with low time overhead and success in achieving small levels of power consumption for operating sensors. Additionally, we implemented the presented approach constructing a small test bed employing nano-enhanced sensing elements. Analyzing such elements showed a clear change in the visible fluorescence intensity with the variation of DO ratio. The reason is the developed oxygen vacancies concentration formed inside ceria nanoparticles which act as DO receptor. The entire system operation is managed by an automated management system. Results showed the clear effect of the smart, trustworthy, and automated data collection and analysis on the quality and accuracy of sensing. Having such system is essential for easy data consolidation and information extraction. Different, experiments were conducted to show the effect of having such automation in enhancing the sensor and the network lifetime even in presence of malicious nodes.

The proposed system is an adequate solution for a comprehensive automated management of DO sensors in the aqueous media. The automation platform within the OMS is built to be generic and can be easily modified to be used with wide variety of other applications and sensing elements achieving wide scope of applications. Our future work includes testing the proposed scheme experimentally for other applications and in large-scale scenarios.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

The presented work is part of the awarded Grant “ARP2013 .R13.2” funded by Information Technology Industry Development Agency (ITIDA Egypt).

References


