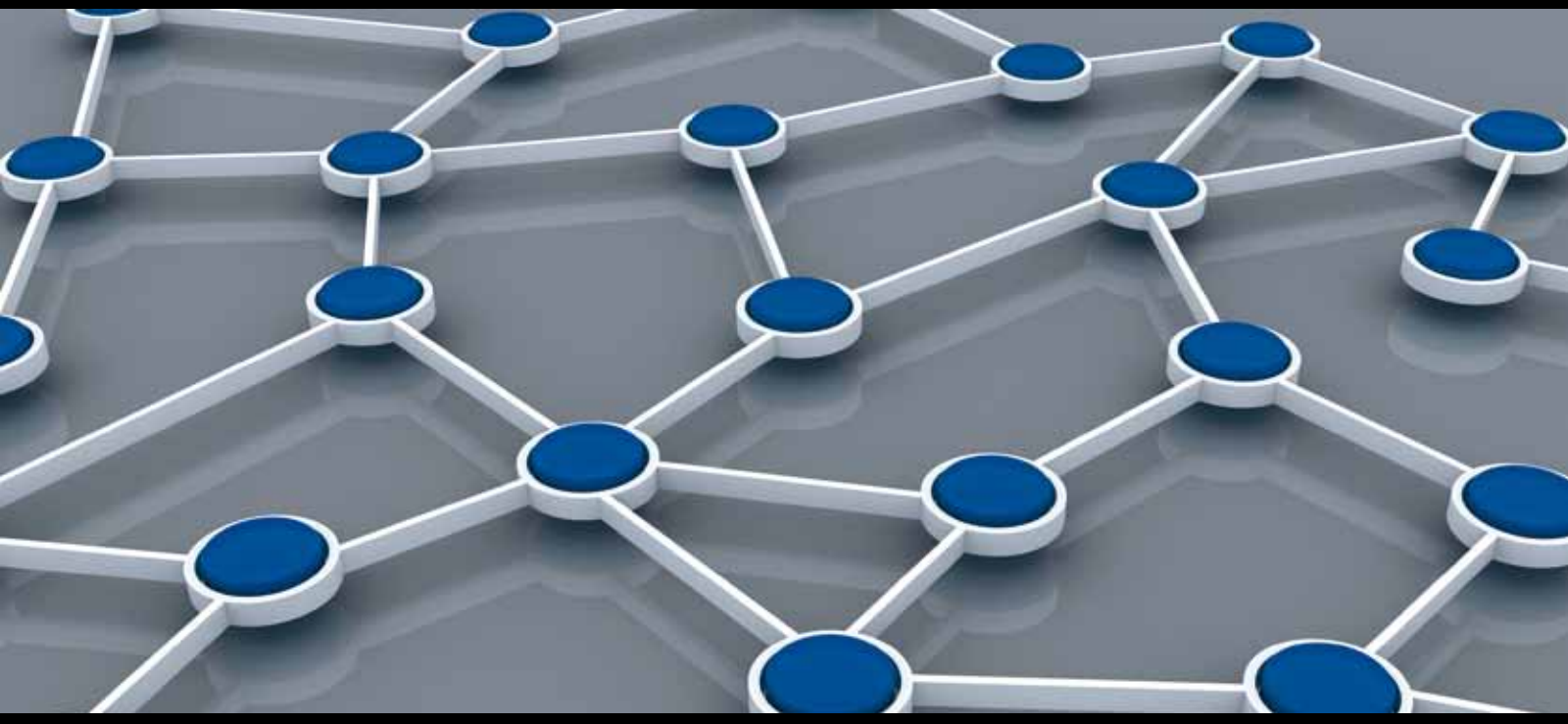
International Journal of Distributed Sensor Networks

Advanced Sensor Technology and Applications in Industrial Control System

Guest Editors: Tai-Hoon Kim, Ruay-Shiung Chang, Carlos Ramos, and Sabah Mohammed



Advanced Sensor Technology and Applications in Industrial Control System

International Journal of Distributed Sensor Networks

Advanced Sensor Technology and Applications in Industrial Control System

Guest Editors: Tai-hoon Kim, Ruay-Shiung Chang, Carlos Ramos, and Sabah Mohammed

Editorial Board

Contents

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 516412, 2 pages
http://dx.doi.org/10.1155/2013/516412

Editorial

Advanced Sensor Technology and Applications in Industrial Control System

Tai-hoon Kim,[1] Ruay-Shiung Chang,[2] Carlos Ramos,[3] and Sabah Mohammed[4]

[1] *GVSA and University of Tasmania, Australia*
[2] *National Dong Hwa University, Taiwan*
[3] *Polytechnic Institute of Porto, Portugal*
[4] *Lakehead University, Canada*

Correspondence should be addressed to Tai-hoon Kim; taihoonn@empal.com

Received 16 September 2013; Accepted 16 September 2013

On behalf of the Science and Engineering Research Support society (SERSC), it is an honor for us to introduce you to this special issue, which includes articles that cover topics of particular interest to researchers in the field of sensor technology and applications in industrial control system.

This issue contains 18 articles that come from various countries, among which we mention Turkey, USA, Macau, Taiwan, Republic of Korea, Spain, China, and Italy. Achieving such a high quality of papers would have been impossible without the huge work that was undertaken by the External Reviewers. We take this opportunity to thank them for their great support and cooperation.

In "*Continuous monitoring of water quality using portable and low-cost approaches*," the authors proposed two different approaches for autonomous monitoring of water quality. In the first system, the boats receive trajectories set by the operator before a mission and follow the trajectories to visit predefined sampling points. In the second system, a group of portable water quality probes mounted on buoys at fixed positions regularly analyze water quality and send the data to the control center.

Authors proposed a method for correction of frequency-sweep nonlinearity in a signal processor instead of linearization of the VCO frequency sweep. For linearization, an additional fixed delay structure was adopted, and the frequency-sweep nonlinearity was extracted and used for compensation in "*Correction of nonlinear frequency sweep in frequency-modulated continuous-wave laser range sensor*." The authors validated their linearization method with numerical analysis and computer simulation.

The main contribution of the paper "*Target localization in wireless sensor networks for industrial control with selected sensors*" was a novel energy-based target localization method in WSNs with selected sensors. In this method, sensors used *turbo product* code (TPC) to transmit decisions to the fusion center. Moreover, the thresholds used in authors' target localization method were determined using a new heuristic method specifically designed for sensor position following uniform distributions and target position following a Gaussian distribution.

In "*Fisher information of mine collapse hole detection based on sensor nodes connectivity*," the authors presented the possibility of detecting the collapse hole using WSN. By establishing a 2D model of the collapse hole in coal mine, the authors described a class of algorithms for detecting the collapse hole in coal mine. Based on log-normal shadowing channel model, the authors analyzed the accuracy of detecting the collapse hole in coal mine using Fisher information and made numerical calculation.

An integrated system for the monitoring of the patient's history must be capable of real-time behavior. Moreover, immediately needed Tag information for monitoring requires a fast-booting system, and one of the essential requirements for mobile conversion is a low-power mobile system. In "*The display system of a patient's history using the RFID and Linux*," to monitor RFID data in real time, the RFID system used a 125 kHz carrier wave with the EM4095 and an embedded Linux operating system with a 400 MHz PXA255 ARM RISC chip, a 512 Kbyte Boot Flash, and a 64 Mbyte SDRAM. Moreover, the system was configured to use a NAND Flash.

In "*Multiple odor recognition and source direction estimation with an electronic nose system*," the authors proposed a mobile olfaction system that was composed of multiple odor recognition and odor source direction estimation abilities for extensive use in mobile environments. Also, authors suggested a recognition algorithm capable of detecting two odors simultaneously using a hierarchical elimination method.

In "*Novel MAC protocol and middleware designs for wearable sensor-based systems for health monitoring*," authors proposed a middleware platform built on WUSB (Wireless USB) over WBAN (Wireless Body Area Networks) hierarchical protocol for wearable health-monitoring systems (WHMS). The proposed middleware platform is composed of time synchronization and localization solutions. And it is executed on the basis of WUSB over WBAN protocol at each sensor node comprising the WHMS.

In "*Optimizing classification decision trees by using weighted naive bayes predictors to reduce the imbalanced class problem in wireless sensor network*," authors devised a new scheme that extends a popular stream classification algorithm to the analysis of WSNs for reducing the adverse effects of the imbalanced class in the data. This new scheme was resource-light at the algorithm level and does not require any data preprocessing. Experiments showed that authors' modified algorithm outperforms the original stream classification algorithm.

In "*On the impact of local processing for motor monitoring systems in industrial environments using wireless sensor networks*," authors presented a theoretical study for verifying the performance of motor monitoring systems in industry employing WSN. First, a discussion was performed about the standards and protocols already proposed for WSN, and about some implementation aspects, which can impact the quality of service in WSN based applications in industrial environments. Finally, mathematical models were developed for verifying the performance of an IEEE 802.15.4 based WSN for applications of torque and efficiency monitoring in induction motors, which are widely used in industries.

In "*Robust people tracking using an adaptive sensor fusion between a laser scanner and video camera*," authors presented an adaptive sensor fusion method between the laser scanner and video camera. In this proposed approach, authors' system does not need a checkerboard. To evaluate the performance of the proposed system, authors showed error analysis and two applications.

In "*Multiple Interface Parallel Approach of Bioinspired Routing Protocol for Mobile Ad Hoc Networks*," authors presented a new bioinspired routing protocol for mobile ad hoc networks obtained thanks to new parallelization techniques of a base protocol called AntOR which has two versions: the so-called disjoint-link (AntOR-DLR) and disjoint-node (AntOR-DNR). The new parallel approach (PAntOR-MI) used the disjoint-node version of AntOR (AntOR-DNR) as the main protocol, as well as the existing (PAntOR).

Presented in "*A PEM fuel cell diagnostic system using the extension theory*" was a fault diagnostic system for a fuel cell, an easy to implement system made up of multiple Modbus modules. On top of that, a user friendly human machine interface was constructed for easy monitoring of the cell system operation. Integrated with a ZigBee wireless communication module, this system can be in the future applied to a distant monitoring system, such as an alter system for a fuel cell powered vehicle.

Because the existing storage scheme for one-dimensional is not suitable for the multi-dimensional data or costs too much energy, in "*Data storage scheme supporting for multi-dimensional query*," authors proposed a kind of data storage scheme supporting multidimensional query inspired by K-D tree. The scheme can effectively store the high-dimensional similar data to the same piece of two-dimensional area. It can quickly fix the storage area of the event by analyzing the query condition and then fetch back the query result.

The main purpose of "*Deadline-aware scheduling perspectives in industrial wireless networks: a comparison between IEEE 802.15.4 and Bluetooth*" is to reduce, as much as possible, the packet loss on the channel, increasing, at the same time, the reliability of the wireless technology. The simulation campaign clearly demonstrated that the combined use of EDF (for periodic traffic flows management) and CBS (per aperiodic traffic flows management) determined more enhancements in IEEE 802.15.4 than Bluetooth.

In "*Security issues in mobile ad hoc networks*," a few existing proposals in the field of secure autoconfiguration in MANETs were presented, and they were examined against seven of the most common threats that can be found on these kind of networks to determine how secure or vulnerable they are.

In "*Immune embedded Linux core system with multiple sensors*," by building an immune mechanism based on the normal model, an immune embedded Linux core system with multiple sensors was presented. In this system, the sensors were used to collect the data concerning the environment information of the disaster.

Authors mainly focused on Bluetooth roaming sensor technology and designed a sensor protocol which works in roaming for Bluetooth multisystems technology in their paper "*Sensor protocol for roaming bluetooth multiagent systems*." The advantage of designing a roaming protocol is to ensure that the Bluetooth-enabled roaming devices can freely move inside the network coverage without losing their connection or break of service in case of changing the base stations.

In "*Wireless virtual multiple antenna networks for critical process control: protocol design and experiments*," the author evaluated experimentally the impact of fading channels on the controllability of the closed-loop wireless system. In particular, it is envisaged here that the incorporation of the cooperative network paradigm into future wireless system standardization will allow cable replacing in tight closed-loop control applications with cycle-time below 100 ms.

Tai-hoon Kim
Ruay-Shiung Chang
Carlos Ramos
Sabah Mohammed

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 306496, 7 pages
http://dx.doi.org/10.1155/2013/306496

Research Article

Fisher Information of Mine Collapse Hole Detection Based on Sensor Nodes Connectivity

Shengbo Hu,[1,2] Heng Shu,[1,2] and Xiaowei Song[1,2]

[1] *Institute of Intelligent Information Processing, Guizhou Normal University, Guiyang 550001, China*
[2] *Center for RFID and WSN Engineering, Department of Education, Guiyang, Guizhou 550001, China*

Correspondence should be addressed to Shengbo Hu; hsb@nssc.ac.cn

Received 15 January 2013; Accepted 14 August 2013

Academic Editor: Tai-hoon Kim

It is very important to detect a collapse hole for coal mine workers. The possibility of detecting the collapse hole using WSN is presented because the tunnel in coal mine is narrow and has poor working condition. Comparing three types of the hole detection methods, it is seen that the connectivity-based methods are used to detect coal mine collapse better than other methods. By establishing a 2D model of the collapse hole in coal mine, a class of algorithms for detecting the collapse hole in coal mine is described. Based on log-normal shadowing channel model, the accuracy of detecting the collapse hole in coal mine using Fisher information is analyzed. Numerical calculation shows that connectivity-based localization schemes are better to detect collapse hole of coal mine.

1. Introduction

Coal mine collapse is one of the main reasons of coal mine fatalities in the past 10 years in the world [1]. Hence, it is very important to detect a collapse hole and accurately provide location references for coal mine workers. Since the coal mine collapse may destroy some coal safety monitoring devices, detecting the collapse hole in coal mine becomes a great challenge. The utilization of wire sensors to monitor coal mine is the primary method at present. However, the wired method makes the monitoring systems less scalable and vulnerable because of poor working conditions in the tunnel of coal mine. Once coal mine collapse happens, all wired sensors may be destroyed, and it is impossible to detect the collapse hole.

Wireless sensor network (WSN) is an event-based self-organized wireless network that relies on deploying spatially dense sensor nodes observing a physical phenomenon [2, 3]. Compared with traditional wire sensing, WSN can achieve larger coverage area, greater accuracy, and more flexible deployment. The utilization of WSN to monitor coal mine is of benefit. Once coal mine collapse happens, not all sensor nodes of WSN are destroyed. So, it is possible to detect the collapse hole in coal mine. For example, Li and Liu present a method based on a regular beacon strategy to detect the collapse hole in coal mine by regulating a mesh sensor network deployment [1].

When a collapse happens, the group of destroyed sensor nodes in the coal mine wireless sensor network creates a hole [4]. The collapse hole boundary separates all the faulty sensor nodes from the working nodes. Generally, there can be three types of the hole detection methods. The approaches based on geometric location methods [5, 6] rely on the nodes having geographical locations and can find more accurate boundary nodes than the other two methods, but each node has to be equipped with an extra device such as GPS to obtain the geographical locations. Unlike above approaches, the statistical methods [7, 8] without location information usually assume the sensor nodes are uniformly distributed on the sensing field. The major weakness of the statistical methods is that the criteria for detecting hole acquired from the statistical characteristics cannot guarantee to find hole precisely. The connectivity-based methods [9–11] use the information of neighboring sensor nodes connectivity to detect the hole. Normally, the method has higher packet control overhead than the previous two methods due to having to collect

information from neighboring sensor nodes. However, it does not need location information and has better accuracy of finding boundary nodes than the statistical method. The sensor nodes neither can be equipped with any additional devices such as GPS nor are uniformly distributed because of the narrow tunnel and the poor working conditions of it in coal mine. So, the connectivity-based methods are used to detect coal mine collapse better.

Connectivity is just a binary variable determined by whether or not a sensor node can demodulate and decode a packet transmitted by another sensor node. Connectivity measurements can be obtained by comparing the *received signal strength* (RSS) value between the two nodes against a power threshold. The RSS can be used to implement range-based localization [12]. The localization methods are popular because no additional hardware is required on the sensor nodes. However, the range estimates using RSS are inaccurate and can lead to large localization error because the RSS value is affected by unpredictable shadowing and fading in the tunnel in coal mine. Yet connectivity is the binary variable carrying information regarding sensor nodes location and is often discussed without considering the effect of shadowing and fading channel. So, the connectivity-based methods for localization have been actively researched in hole detection and ad hoc routing in WSN [13–15].

Connectivity measurements are actually just a binary quantization of the RSS measurements against a power threshold. Because RSS measurements are affected by unpredictable shadowing and fading in the tunnel in coal mine, it is very important to analyze accuracy of detecting the coal mine collapse hole using the connectivity-based methods. The rest of this paper is organized as follows. Section 2 discusses the system models including a hierarchy WSN topology in coal mine underground and 2D model of the coal mine collapse hole. Section 3 introduces a class of algorithms for detecting the coal mine collapse hole using the connectivity-based methods. Section 4 analyzes the accuracy of detecting collapse hole using Fisher information. Section 5 presents numerical calculation and discussion. Section 6 concludes this paper.

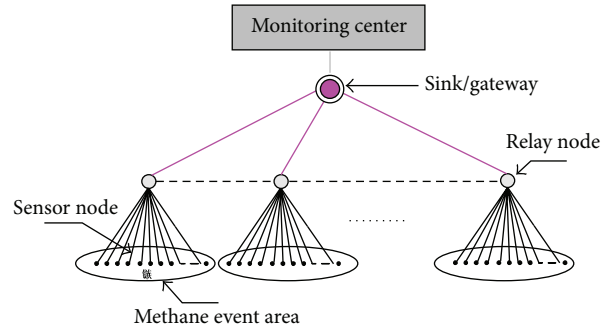


Figure 1: Hierarchy of nodes in an underground WSN in coal mine.

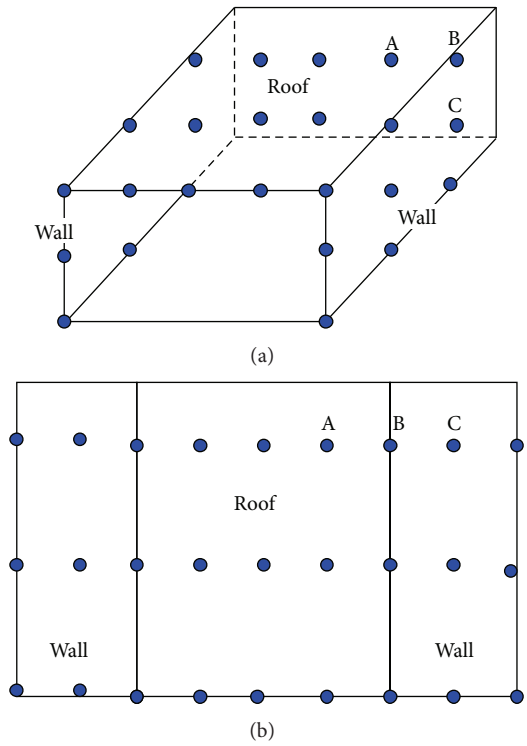


Figure 2: A cluster of sensor nodes deployment.

2. System Models

2.1. A Hierarchy Underground WSN Topology in Coal Mine. In view of the characteristics of narrow and poor environment in coal mine underground, it is better to use the hierarchy WSN topology for improving coverage and connectivity [16, 17]. The hierarchy WSN comprises sensor nodes, relay nodes, and sink nodes. The sensor nodes gather methane data autonomously. Multisensors form a cluster, and the cluster heads are the addresses of relay nodes. Methane data collected by the sensor nodes are forwarded to sink nodes through relay nodes. Communication between relay nodes forms a multihop route, and monitoring data are processed at sink nodes. Finally, the gateway passes the data processed at sink nodes to methane monitoring center. Table 1 shows the various types of nodes in the hierarchy WSN. These nodes will form a hierarchy that is shown in Figure 1.

Table 1: Types of nodes in hierarchy WSN in coal mine underground.

Type	Energy	Location	Role
Sensor node	Constrained	Random	Sensing
Relay node	Not constrained	Fixed	Forwarding
Sink node	Not constrained	Fixed	Processing

2.2. 2D Model of Coal Mine Collapse Hole. Generally, the tunnel of coal mine can be classified as arch-shaped, rectangle, trapezium, or semicircle tunnel. For convenience, we consider the rectangle tunnel in this paper. A cluster of sensor nodes are deployed on the walls and roofs of the tunnel, as shown in Figure 2(a). To facilitate collapse hole detection, Figure 2(a) can be unfolded to a 2D representation as depicted in Figure 2(b).

Comparing Figure 2(a) with Figure 2(b), the relationships between the neighboring sensor nodes in Figure 2(b) are the same as in Figure 2(a). However, the distance between any two nodes in Figure 2(b) is greater than or equal to the distance between the pair in Figure 2(a). Thus, the real connectivity of the WSN is no less than shown in the 2-D representation in Figure 2(b), and the accuracy of the collapse hole detection in Figure 2(b) is preserved in Figure 2(a).

So, a cluster of sensor nodes are modeled as a 2D graph, $G = (V, E)$, where each vertex represents a sensor node, V is the set of a cluster of sensor nodes, and two vertices are connected by an edge in E if and only if their distance is at most the guaranteed communication radius. Once a mine collapse occurs, some sensor nodes $n_1, n_2, \ldots, n_k \in V$ are destroyed. The extent of these damaged nodes creates a collapse hole with convex hulls, which is surrounded by alive sensor nodes that contain all the damaged sensor nodes in Figure 3.

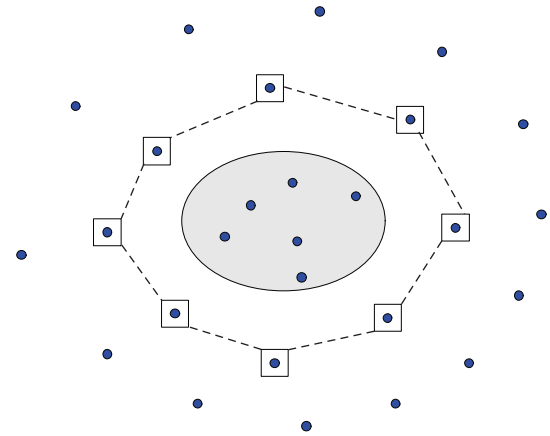


Figure 3: A collapse hole created by damaged sensor nodes.

3. Algorithm for Detecting Collapse Hole Using Connectivity

Connectivity simply reports whether or not a sensor node can demodulate and decode a packet transmitted by another sensor node. Connectivity measurements can be obtained by comparing the *received signal strength* (RSS) value between the two nodes against a power threshold. Because the RSS measurements are affected by unpredictable shadowing and fading in the tunnel in coal mine, the connectivity can be described by a random binary variable Q_{ij}:

$$Q_{ij} = \begin{cases} 1, & P_{ij} \geq P_{\text{th}}, \\ 0, & P_{ij} < P_{\text{th}}, \end{cases} \tag{1}$$

where P_{ij} is the RSS measurement received at sensor node i transmitted by sensor node j and P_{th} is a power threshold of sensor node. Equation (1) shows that sensor node i can demodulate and decode a packet transmitted by sensor node j if $P_{ij} \geq P_{\text{th}}$, and $Q_{ij} = 1$; sensor node i cannot demodulate and decode a packet transmitted by sensor node j if $P_{ij} < P_{\text{th}}$, and $Q_{ij} = 0$.

The connectivity-based methods will use the assumption that two sensor nodes are "connected" if $Q_{ij} = 1$ and "disconnected" if $Q_{ij} = 0$. The advantage of these methods is that location can be discussed without knowing the propagation model's parameters. So, a class of algorithms for detecting the collapse hole in coal mine can be described as follows [1, 18].

Step 1 (initialization). Each sensor node broadcasts a ping requesting information to neighboring sensor nodes. If $P_{ij} \geq P_{\text{th}}$, sensor nodes reply with their ID, and the pinging node can create a list of its neighboring sensor nodes.

Step 2 (integrity checking). If a coal mine collapse occurs, each sensor node pings its neighboring sensor nodes, keeps track of the responses, and compares the list of neighboring sensor nodes. If $P_{ij} \geq P_{\text{th}}$, sensor node i receives the ID (identification number) of sensor node j and marks sensor node j as alive. If $P_{ij} < P_{\text{th}}$, sensor node i cannot receive ID of sensor node j and marks the sensor node as missed. If the node's number of missed neighboring nodes exceeds a threshold, sensor node i marks itself as belonging to the collapse perimeter.

Step 3 (collapse hole scanning). According to the list of sensor nodes marked as "on the collapse perimeter," the classical Graham algorithm [19] is used to detect the collapse hole with convex hulls.

4. Accuracy of Detecting Collapse Hole

As shown in Section 3, analyzing the accuracy of detecting collapse hole using connectivity is equivalent to analyzing the accuracy of estimating distance d between neighboring sensor nodes equivalently using Q_{ij}. So, we consider the topic of estimate distance d between neighboring sensor nodes from connectivity measurements using Fisher information.

4.1. Channel Model. Because the tunnel is narrow and the environment is poor in coal mine, the RSS is attenuated by path losses, fading, and shadowing losses [20]. Path loss is the deterministic reduction function of distance d between neighboring sensor nodes. Fading is the effect of multipath propagation. Because many wireless sensor nodes use spread-spectrum techniques, the fading can be reduced mostly and its impact on the attenuated RSS is not significant. Shadowing is the loss incurred as a signal passes through permanent obstructions (e.g., walls or buildings). For most sensor nodes in the tunnel, shadowing losses cannot be countered. When a collapse occurs, shadowing losses are greater. So, the RSS values follow the log-normal shadowing model, a channel model widely used in WSN [21, 22].

Let $P_r(0)$ be the received signal power at sensor node with the reference distance d_0 (typically $d_0 = 1$) and let γ be the path loss exponent, a parameter that depends on the environment where communication occurs (typical values are 2 and 4 [23]). In log-normal shadowing model, the

received power when the two sensor nodes are at a distance d can be written as

$$\ln P_r(d) = \ln P_r(0) - \gamma \ln d + w, \tag{2}$$

where w is a zero-mean random variable with normal distribution $N(0, \sigma^2)$.

4.2. Fisher Information. In this paper, we focus on estimating distance d between neighboring sensor nodes from connectivity measurements, which carry information regarding sensor nodes location. As is well known, the Fisher information measures the amount of information that a random variable carries about an unknown parameter. And the inverse of the Fisher information, known as the Cramer-Rao Bound, is the minimum variance for any unbiased estimator. Here, the random variable Q_{ij} defined by (1) is used to estimate d. So, $\mathrm{Var}\{\widehat{d}\}$, the variance of estimating d, is

$$\mathrm{Var}\left\{\widehat{d}\right\} \geq \frac{1}{I(d)}, \tag{3}$$

where $I(d)$ is the Fisher information of estimating distance d using the random variable Q_{ij}.

4.2.1. Fisher Information of RSS Measurements. For notational convenience, we define

$$\alpha = \frac{2}{\gamma}, \qquad z = \left(\frac{P_r(0)}{P_r(d)}\right)^{\alpha}. \tag{4}$$

Then, (2) can be written as

$$z = d^2 e^{-\alpha w}. \tag{5}$$

So, the underlying estimation problem is to estimate d from the RSS measurements z satisfying (5), given the knowledge of α and σ^2.

Taking the logarithm of (5), we obtain

$$\ln z = 2 \ln d - \alpha w. \tag{6}$$

Let $l = \ln z$, where l is a random variable with normal distribution $N(2 \ln d, \alpha^2 \sigma^2)$.

The log-likelihood function is given by

$$\ln\left[p(l, y)\right] = -\ln\left[2\pi\alpha^2\sigma^2\right] - \frac{(l - 2\ln d)^2}{2\alpha^2\sigma^2}. \tag{7}$$

Taking the derivative of the log-likelihood function, we obtain

$$\frac{\partial \ln\left[p(l, y)\right]}{\partial d} = \frac{2(l - 2\ln d)}{d\alpha^2\sigma^2}. \tag{8}$$

Hence, the Fisher information of RSS measurements is given by

$$I_{\mathrm{RSS}}(d) = E\left[\left(\frac{\partial \ln\left[p(l, y)\right]}{\partial d}\right)^2\right] = \frac{4}{d^2\alpha^2\sigma^2} = \frac{\gamma^2}{d^2\sigma^2} = \frac{k^2}{d^2}, \tag{9}$$

where $k = \gamma/\sigma$.

4.2.2. Fisher Information of Connectivity Measurements. As shown in Section 3, the underlying estimation problem is to estimate d from the connectivity measurements. Then the Fisher information of connectivity measurements depends not only on d between neighboring sensor nodes, but also on the value of the power threshold P_{th} in (1). For convenience, P_{th} needs to be converted into the distance threshold d_{th}. From (5), it can be observed that l has a nonaffine dependence on d and an affine dependence on w. Hence, no efficient estimator exists for this problem [24]. This leads us to choose the maximum likelihood estimator (MLE) $\widehat{d}_{\mathrm{th}}$ of the distance threshold d_{th}. The $\widehat{d}_{\mathrm{th}}$ is given by

$$\widehat{d}_{\mathrm{th}} = \arg\max_{d}^{P_r(d)=P_{\mathrm{th}}} \ln\left[p(l, y)\right] = \sqrt{\left(\frac{P_r(0)}{P_{\mathrm{th}}}\right)^{\alpha}} = \left(\frac{P_r(0)}{P_{\mathrm{th}}}\right)^{1/\gamma}. \tag{10}$$

From [25], consider the case of 2-level quantized RSS. The Fisher information of connectivity measurements is given by

$$\begin{aligned} I_{\mathrm{CON}}(d, d_{\mathrm{th}}) &= \frac{4}{d^2\alpha^2\sigma^2} h_r(d, d_{\mathrm{th}}) \\ &= \frac{\gamma^2}{d^2\sigma^2} h_r(d, d_{\mathrm{th}}) = \frac{k^2}{d^2} h_r(d, d_{\mathrm{th}}), \end{aligned} \tag{11}$$

where the term $h_r(d, d_{\mathrm{th}})$ depends on the ratio between d and d_{th}:

$$h_r(d, d_{\mathrm{th}}) = \frac{2}{\pi} \frac{\exp\left[-k^2 \ln\left(d/d_{\mathrm{th}}\right)^2\right]}{1 - \mathrm{erf}\left[k \ln\left(d/d_{\mathrm{th}}\right)/\sqrt{2}\right]^2}, \tag{12}$$

where $\mathrm{erf}(\cdot)$ is an error function.

5. Numerical Calculation and Discussion

To analyze the accuracy of detecting collapse hole based on connectivity measurements, we examine the Fisher information using numerical calculation in this section.

5.1. Numerical Calculation. In this subsection, we investigate the Fisher information $I_{\mathrm{RSS}}(d)$ and $I_{\mathrm{CON}}(d, d_{\mathrm{th}})$ using numerical calculation.

5.1.1. The Path Loss Exponent's Effects on the Fisher Information $I_{RSS}(d)$. Figure 4 shows that $I_{\mathrm{RSS}}(d)$ is a function of d for different γ values ($\gamma = 4, 3, 2.5$) and fixed σ value ($\sigma = 2$).

Figure 4 describes that the amount of the $I_{\mathrm{RSS}}(d)$ available to estimate d decreases for increasing values of the distance d and increases for increasing values of γ. This shows that the estimates become more accurate because the attenuation caused by the path loss clears the RSS measurement.

5.1.2. The σ Values' Effects on the Fisher Information $I_{RSS}(d)$. Figure 5 shows that $I_{\mathrm{RSS}}(d)$ is a function of d for different σ values ($\sigma = 2, 3, 4$) and fixed γ value ($\gamma = 3$).

Figure 5 describes that the amount of $I_{\mathrm{RSS}}(d)$ available to estimate d decreases for increasing values of the distance

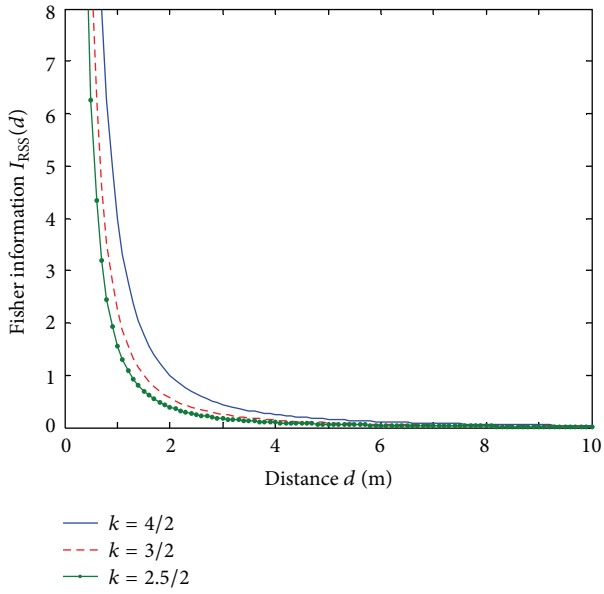


FIGURE 4: Fisher information for RSS measurements at fixed σ value.

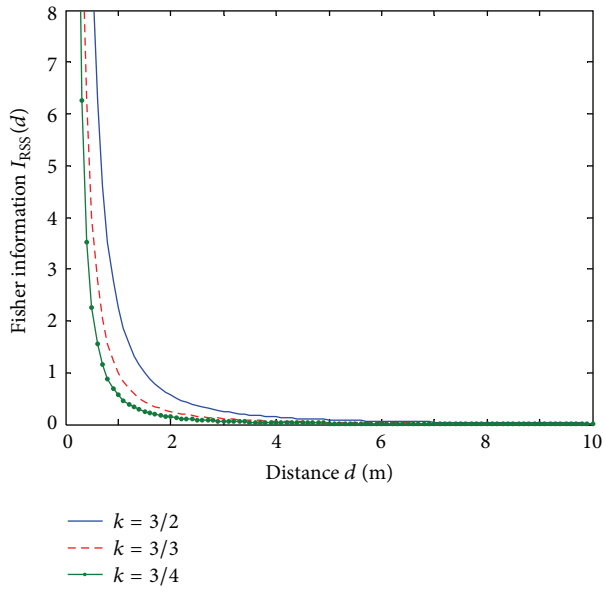


FIGURE 5: Fisher information for RSS measurements at a fixed γ value.

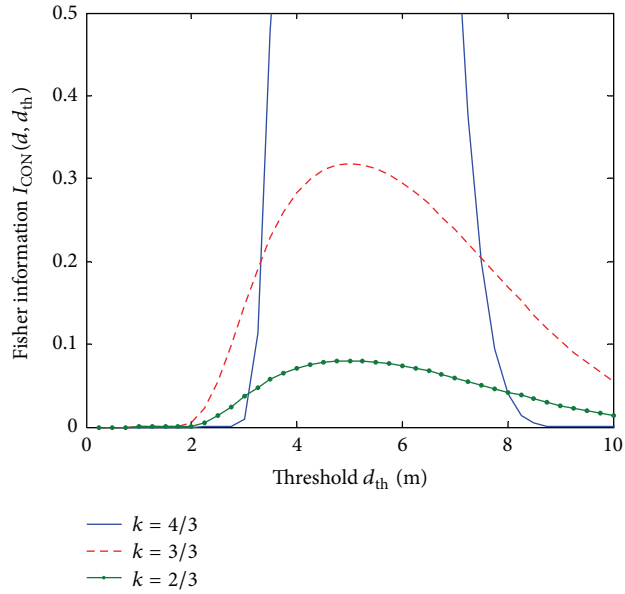


FIGURE 6: Fisher information for connectivity measurements at different k values.

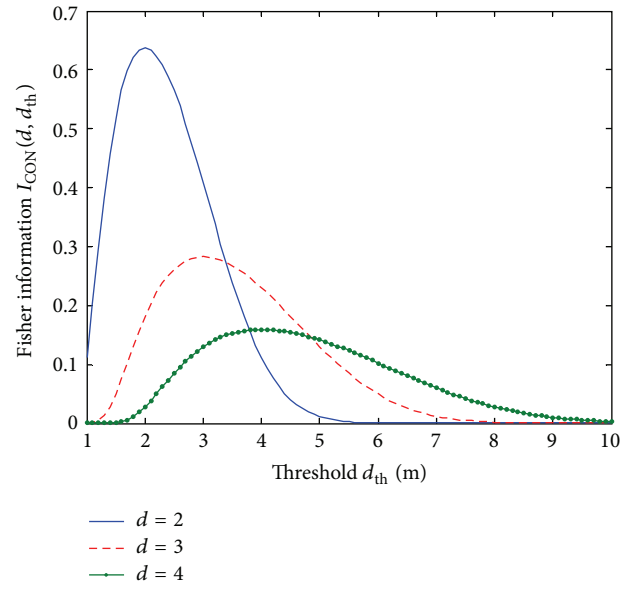


FIGURE 7: Fisher information for connectivity measurements at different d values.

d and increasing values of σ. This shows that the estimates become less accurate because the variability caused by the RF shadowing blurs the RSS measurement.

5.1.3. The k Values' Effects on the Fisher Information $I_{CON}(d,d_{th})$. Figure 6 shows that $I_{CON}(d, d_{th})$ is a function of d_{th} when $d = 5$ m for different k values ($k = 3/4, 3/3, 3/2$), and Figure 7 shows that $I_{CON}(d, d_{th})$ is a function of d_{th} when $k = 4/2$ for different d values ($d = 2, 3, 4$).

Figures 6 and 7 describe that $I_{CON}(d, d_{th})$ always peaks when $d_{th} = d$. In other words, connectivity measurements reach the maximum information if the distance threshold is equal to the true distance between neighboring sensor nodes. From (12), the maximum Fisher information of connectivity measurements $I_{CON}(d, d_{th}) = 0.63 I_{RSS}(d)$ if $d_{th} = d$. So, $I_{CON}(d, d_{th})$ is always lower than $I_{RSS}(d)$.

5.2. Discussion about Accuracy of Detecting Collapse Hole. As shown in Section 5.1, RSS measurements carry greater Fisher information than connectivity ones. However, this is only true as long as the neighboring sensor nodes are within the radio coverage of each other. When sensor nodes are within

each other's radio coverage, they are able to communicate and successfully exchange radio message.

Depending on the choice of P_{th}, two neighboring sensor nodes that are within each other's radio range can be considered connected or disconnected. When the neighboring sensor nodes are far from radio coverage, they will not be able to communicate. So, no RSS information can be collected, and $I_{\text{RSS}}(d) = 0$. Instead, the sensor nodes far from radio coverage can be associated with the value Q_{ij}, and $I_{\text{CON}}(d, d_{\text{th}}) > 0$.

As shown above, RSS range-based localization schemes are more accurate when the neighboring sensor nodes are in the radio coverage of each other, but connectivity-based localization schemes are naturally suited to localize nodes that are far from radio coverage. Hence, connectivity-based localization schemes are used to detect collapse hole of coal mine better.

6. Conclusion

It is very important to detect a collapse hole for coal mine workers. This paper presents the possibility of detecting the collapse hole using WSN. Comparing three types of hole detection methods, we think that the connectivity-based methods are better to detect coal mine collapse. By establishing a 2D model of the collapse hole in coal mine, we describe a class of algorithms for detecting the collapse hole in coal mine. Based on log-normal shadowing channel model, we analyze the accuracy of detecting the collapse hole in coal mine using Fisher information and make numerical calculation. Finally, we make the conclusion as follows.

(i) The larger the path loss exponent, the greater the Fisher information.

(ii) The larger the variability caused by the RF shadowing, the less the Fisher information.

(iii) When sensor nodes are within each other's radio coverage, the accuracy of detecting the collapse hole *using RSS measurements is higher than that when using connectivity ones*; when the neighboring sensor nodes are far from radio coverage, no RSS information can be collected, and $I_{\text{RSS}}(d) = 0$. Instead, the sensor nodes far from radio coverage can be associated with the value Q_{ij}, and $I_{\text{CON}}(d, d_{\text{th}}) > 0$. Hence, connectivity-based localization schemes are used to detect collapse hole of coal mine better.

Acknowledgments

The authors wish to thank the editor and reviewers for their valuable comments, corrections, and suggestions, which led to an improved version of the original paper. This research is a project partially supported by the National Natural Science Foundation of China (Grant no. 61064614), Guizhou Science and Technology Innovation Group for RFID & WSN, and Guizhou Natural Science Foundation (Grant no. [2012] 38).

References

[1] M. Li and Y. Liu, "Underground coal mine monitoring with wireless sensor networks," *ACM Transactions on Sensor Networks*, vol. 5, no. 2, article 10, 2009.

[2] K. Kavi, P. Rajiv, and M. Avinash, "A wireless sensor network air pollution monitoring system," *International Journal of Wireless and Mobile Networks*, vol. 12, no. 2, pp. 31–45, 2010.

[3] W. U. Bajwa, A. Sayeed, and R. Nowak, "Matched source-channel communication for field estimation in wireless sensor networks," in *Proceedings of the 4th International Symposium on Information Processing in Sensor Networks (IPSN '05)*, pp. 332–339, Los Angeles, Calif, USA, April 2005.

[4] Y. Wang, J. Gao, and J. S. B. Mitchell, "Boundary recognition in sensor networks by topological methods," in *Proceedings of the 12th Annual International Conference on Mobile Computing and Networking (MOBICOM '06)*, pp. 122–133, Los Angeles, Calif, USA, September 2006.

[5] Q. Fang, J. Gao, and L. J. Guibas, "Locating and bypassing holes in sensor networks," *Mobile Networks and Applications*, vol. 11, no. 2, pp. 187–200, 2006.

[6] P. K. Sahoo, K.-Y. Hsieh, and J.-P. Sheu, "Boundary node selection and target detection in wireless sensor network," in *Proceedings of the 4th IEEE and IFIP International Conference on Wireless and Optical Communications Networks (WOCN '07)*, Singapore, July 2007.

[7] S. P. Fekete, M. Kaufmann, A. Kroller, and N. Lehmann, "A new approach for boundary recognition in geometric sensor networks," in *Proceedings of 17th Canadian Conference on Computational Geometry*, Prince Edward Island, Canada, August, 2005.

[8] S. P. Fekete, A. Kröller, D. Pfisterer, S. Fischer, and C. Buschmann, "Neighborhood-based topology recognition in sensor networks," in *Proceedings of 1st International Workshop on Algorithmic Aspects of Wireless Sensor Networks*, Turku, Finland, July 2004.

[9] S. Funke and C. Klein, "Hole detection or: 'How much geometry hides in connectivity?'," in *Proceedings of the 22nd ACM Annual Symposium on Computational Geometry (SCG '06)*, pp. 377–385, Sedona, Ariz, USA, June 2006.

[10] O. Saukh, R. Sauter, M. Gauger, P. J. Marrón, and K. Rothermel, "On boundary recognition without location information in wireless sensor networks," in *Proceedings of the 7th International Conference on Information Processing in Sensor Networks (IPSN '08)*, pp. 207–218, St. Louis, Mo, USA, April 2008.

[11] Y. Wang, J. Gao, and J. S. B. Mitchell, "Boundary recognition in sensor networks by topological methods," in *Proceedings of the 12th Annual International Conference on Mobile Computing and Networking (MobiCom '06)*, Los Angeles, Calif, USA, September 2006.

[12] A. Savvides, C. Han, and M. B. Strivastava, "Dynamic fine-grained localization in ad-hoc networks of sensors," in *Proceedings of the 7th Annual International Conference on Mobile Computing and Networking*, pp. 166–179, July 2001.

[13] S. Babaie1 and S. S. Pirahesh, "Hole detection for increasing coverage in wireless sensor network using triangular structure," *International Journal of Computer Science*, vol. 9, no. 2, pp. 213–218, 2012.

[14] C. F. Huang and Y. C. Tseng, "A survey of solutions to the coverage problems in wireless sensor networks," *Journal of Internet Technology*, vol. 6, no. 1, pp. 1–8, 2005.

[15] R. Nagpal, H. Shrobe, and J. Bachrach, "Organizing a global coordinate system from local information on an ad hoc sensor network," in *Proceedings of 2nd International Workshop (IPSN '03)*, pp. 333–348, Palo Alto, Calif, USA, April 2003.

[16] J. P. Sun, L. Tang, W. Chen, B. Zhang, N. Zhu, and X.-Y. Zhang, "No blind zone placement of methane sensor based on monitoring coverage," *Journal of the China Coal Society*, vol. 33, no. 5, pp. 946–950, 2008.

[17] W. Yang, S. Y. Zhou, and H. Qiao, "Node localization in wireless sensor networks for coal mine security monitoring," *Journal of the China Coal Society*, vol. 32, no. 6, pp. 652–656, 2007.

[18] I. F. Khan and M. Y. Javed, "Hole healing energy aware algorithm for wireless sensor networks," *International Journal of Computer Theory and Engineering*, vol. 2, no. 5, pp. 706–711, 2010.

[19] R. L. Graham, "An efficient algorith for determining the convex hull of a finite planar set," *Information Processing Letters*, vol. 1, no. 4, pp. 132–133, 1972.

[20] T. Rappaport, *Wireless Communication: Principles and Practice*, IEEE Press Piscataway, New York, NY, USA, 1996.

[21] A. Ghasemi and E. S. Sousa, "Asymptotic performance of collaborative spectrum sensing under correlated log-normal shadowing," *IEEE Communications Letters*, vol. 11, no. 1, pp. 34–36, 2007.

[22] P. Stuedi and G. Alonso, "Log-normal shadowing meets SINR: a numerical study of capacity in wireless networks," in *Proceedings of the 4th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON '07)*, pp. 550–559, San Diego, Calif, USA, June 2007.

[23] K. Sohrabi and G. J. Pottie, "Near ground wideband channel measurement in 800–1000 MHz," in *Proceedings of the 50th IEEE VTS Vehicular Technology Conference*, pp. 571–575, Houston, Tex, USA, May 1999.

[24] A. H. Madsen, "On the existence of efficient estimators," *IEEE Transactions on Signal Processing*, vol. 48, no. 11, pp. 3028–3031, 2000.

[25] N. Patwari and A. Hero III, "Using proximity and quantized RSS for sensor localization in wireless networks," in *Proceedings of the 2nd ACM International Workshop on Wireless Sensor Networks and Applications (WSNA '03)*, pp. 20–29, San Diego, Calif, USA, September 2003.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 361378, 7 pages
http://dx.doi.org/10.1155/2013/361378

Research Article

Multiple Odor Recognition and Source Direction Estimation with an Electronic Nose System

Hyeong-Joon Kwon, Dong-Gyu Kim, and Kwang-Seok Hong

College of Information and Communication Engineering, Sungkyunkwan University, 300 Chunchun-dong, Jangan-Gu, Suwon, Kyungki-do 440-746, Republic of Korea

Correspondence should be addressed to Hyeong-Joon Kwon; katsyuki@skku.edu

Received 10 January 2013; Accepted 7 July 2013

Academic Editor: Tai-hoon Kim

We propose an electronic nose system that can perform real time direction estimation of an odor source and multiple odors recognition based on a stereo sensor array for extensive use in mobile environments. The proposed system consists of the following: (1) a method to obtain odor signals using a twin-sensor array, which consists of 16-channel metal oxide semiconductor sensors; (2) a method to estimate the direction of an odor source by analyzing the signal amplitude of each channel in the stereo sensor array; and (3) a method to recognize two odors simultaneously using a hierarchical elimination method. We determine the accuracy of the direction estimation of odor sources and the odor recognition rate in order to verify the performance of the multiple odors recognition method. As a result, we confirm the high estimation performance of the model for the front three-way directions, with a recognition rate of approximately two odors simultaneously.

1. Introduction

Recent studies on cognitive or recognition ability have been based on the viewpoint of human-computer interaction (HCI) by five-sense information processing, which includes vision, hearing, tactile, palate, and olfaction. Related research requires a convergence of mechanical, electronic, and chemical engineering research [1].

Two human abilities are related to olfaction: smelling ability and odor source localization. In the existing research on smelling ability, either pattern recognition or correlation analysis among sensors has been used to recognize various odors. The pattern recognition-based method is based on the formulaic form of an odor signal [2, 3] and extracts an odor pattern from the odor signal. The traditional method in pattern recognition was considered to be a feature extraction method [4]. The correlation analysis-based method identifies the mutual relationships between odor substances and sensors, and it uses these relationships to recognize the odor [5, 6]. This approach calculates the correlation coefficients among the signals of each sensor in the sensor array and then determines the optimal sensor for the target odor. Both of the aforementioned approaches can recognize one odor at a time. In order to recognize two odors at a time, we must consider an additional feature separation method. The current research focuses on odor source localization with a singular chemical sensor. A single sensor detects a single smell of a chemical substance. Because an odor is composed of various chemical substances, an electronic nose that uses only a singular chemical sensor would have difficulties finding an odor that is generated by a mix of various chemical substances [7, 8]. Therefore, the existing electronic nose cannot be used for extensive use in mobile environments.

In this paper, we propose a mobile olfaction system that is composed of multiple odor recognition and odor source direction estimation abilities for extensive use in mobile environments. The proposed system recognizes two odors at a time and estimates the direction of an odor source using signal amplitude. The proposed system is based on a 16-channel twin-sensor array, similar to human capabilities. Namely, the proposed odor processing system consists of 32 channels. Also, we suggest a recognition algorithm capable of detecting two odors simultaneously using a hierarchical elimination method. This paper contributes to various study

and industry fields that utilize a roll-around humanoid robot with mechanical olfaction.

This paper is organized as follows. In Section 2, we describe the previous studies concerning electronic noses in various study fields. In Section 3, we propose an olfaction system using a twin MOS sensor array. In Section 4, we show the recognition rate of two odors simultaneously and the estimation results of the direction of an odor source. In Section 5, we summarize the results of this study and suggest topics for future work.

2. Related Works

An electronic nose consists of mechanical and logical parts. From a mechanical point of view for developing an electronic nose, an electronic nose that has been developed in order to study mechanical olfaction uses a sensor array containing multiple sensors which can sense odorant molecules. The types of odor sensors that have been considered in previous research include Conducting Polymer (CP), Surface Acoustic Wave (SAW), Quartz Crystal Microbalance (QCM), and Optical and Metal Oxide Semiconductor (MOS) sensors [2]. The type of sensor used varies according to the types of target smell, for example, food monitoring, bioprocesses, medical diagnostics, and beverage manufacturing [9, 10].

From a logical point of view for developing an electronic nose, Linear Discrimination Analysis (LDA), Independent Component Analysis (ICA), Hierarchical Cluster Analysis (HCA), Principal Component Analysis (PCA), Artificial Neural Network (ANN), Support Vector Machine (SVM), statistical Analysis of Variance (ANOVA), and nonhierarchical clustering algorithms such as k-means have been used as feature extraction and classification methods for odor recognition [11–17]. A feature extraction stage of an odor by these algorithms uses the representative components of the input odor. The features are used as a reference model when other odors are recognized.

Recently, in a study related to MOS sensor, Lei Zhang et al. have developed an electronic nose for concentration estimation of formaldehyde [15]. It uses ANN ensemble and self-calibration model for implementation of logical part, and they used four MOS sensors and two auxiliary sensors for the temperature and humidity. A 12-bit analog-digital converter is used as interface between the Field Programmable Gate Array (FPGA) processor and the sensors. FPGA can be used for data collection, storage and processing. The logical part of this nose system is connected to a PC. Brudzewski et al. have proposed a method for recognizing the smell of coffee using a differential electronic nose containing an MOS sensor array [16]. Their study showed that a differential electronic nose, applying the special procedure of signal processing, is of sufficient sensitivity for the recognition of the forgery of coffee and performs much better than the classical electronic nose. It uses 12 MOS sensors of Figaro serious, temperature sensor, and humidity sensor. They also developed a feature extraction method using PCA based on an odor signal database and showed recognition and classification using SVM. Wang et al. also developed an electronic nose using PCA and SVM to predict the total viable counts of odor in chilled pork [17]. In particular, they considered the alpha MOS sensor array system of France and three sensor chambers for controlling the temperature, and an 18-MOS sensor was used.

Furthermore, a recent study considered temperature and humidity based on a sensor array using many MOS sensors. The concept of a chamber was introduced, and an odor recognition method for various industry fields has been proposed. However, studies on the logical part and improvement of electronic ability for developing an electronic nose are fairly scarce because the electronic nose has only been examined in the chemistry and mechanical engineering study area. We also studied an electronic nose using an MOS sensor array. Furthermore, in our previous works, we proposed an optimal MOS sensor decision method [4] and a floral scent recognition method using ICA with a correlation coefficient and multiple odor detection [5]. In particular, reference [5] was presented as an interim result for this study. In this current paper, we will present olfaction using a state-of-the-art electronic nose by improving the logical part as discussed in the following section.

3. Proposed Olfaction System Using Stereo Sensor Array

The proposed olfaction system has multiple odor recognition and odor source localization abilities based on a stereo sensor array that is similar to the human olfactory organ. While in previous works the electronic nose was based on a physical theory, the proposed olfaction system is an intelligent electronic nose that emulates human ability. The most significant difference between previous works and this study is that here we propose a way to improve the logical processing ability of the odor signal by using a stereo sensor array similarly to the human olfactory system.

The signal form of multiple odors is composed of a combination of single odor signals, as shown in Figure 1(a). Figure 1(b) shows the waveforms of a signal of multiple odors. The proposed system recognizes two odors simultaneously from an input signal. A sensor shows a formulaic pattern for an odor. When an odor is inputted, the signal of each sensor shows a pattern [11]. The signal of each sensor increases rapidly according to the input odor and then decreases slowly until the odor has disappeared. Studies on mechanical odors using such signals are classified into one of two categories. The first study category is odor recognition, in which the first step involves detecting the odor activity from the entire set of sensor signals. The sample range of odor activity is roughly 200 to 900, in which case the features of an odor activity can be extracted using pattern recognition.

Figure 2 shows a block diagram of the proposed multiple odors recognition method. The system consists of the acquisition of odor data, the detection of odor activity, the extraction of features, and the recognition of odors. Odor data is acquired through a 16-channel sensor array we developed, and the odor is then converted into odor data. In the proposed system, the odor is detected using entropy from the measured

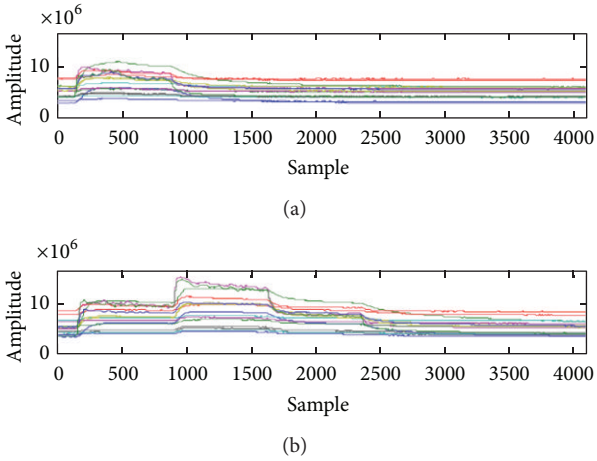


Figure 1: Odor signal waveforms.

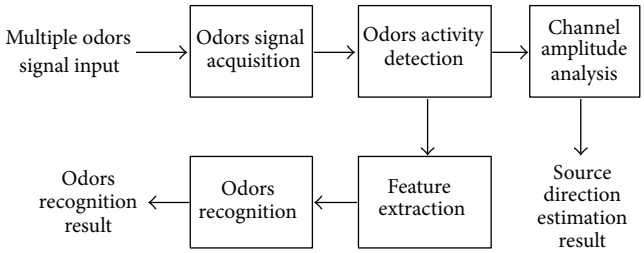


Figure 2: Block diagram of proposed olfaction system.

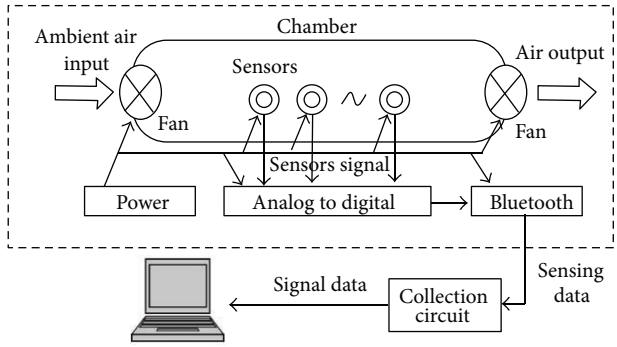


Figure 3: Structure of a sensor array.

Figure 4: Actual image of a stereo sensor array for odor data acquisition.

odor signals, based on the detection of the inflection points of the odor. Features are extracted using a PCA to extract the feature vector of an odor pattern. Subsequently, the odor is recognized using a hierarchical elimination method to select the recognition candidate and by using the Euclidean distance.

An odor signal can be obtained using several different types of sensors for dynamically changing odors in real time. In order to obtain such data, a structure was designed and manufactured that enables the surrounding air to pass through it in real time by attaching fans in the front and back of a chamber. This structure is then used to inhale and discharge the surrounding air into and out of the chamber. The air contains chemicals that constitute a specific odor. The structure of an odor data acquisition device is shown in Figure 3, and an actual image of an odor data acquisition device is shown in Figure 4. As seen in Figure 4, the proposed stereo sensor array was developed based on a socket to easily change a sensor according to the type of target odor.

The odor was detected using entropy in IEEE 802.15 bluetooth-enabled mobile devices (hand-held PC, smartphone, tablet PC, etc.). The odor signals detection module in the logical part of the proposed olfaction system can be used to determine the inflection point and to monitor the abrupt changes in the sensor amplitude during the odor input and removal [4, 11]. In order to detect the inflection point of multiple odors using entropy, a sampling frequency per channel of 13.5 Hz was used. Thirty-two samples per frame were used, with an overlap of 16 samples, and each frame was associated with a calculated entropy value. This information was used to determine the number of sensors with entropies exceeding 1, the condition that occurs when there are more than seven sensors with entropies exceeding 1, and the continuous condition for when there are more than four frames for entropy. A sensor response to prevent incorrect extraction is also included. By adding continuous frame conditions, the abrupt rise in entropy values due to the noise between the actual inflection point and the calculated inflection point was eliminated, and thus incorrect extraction could be prevented. A total of 256 samples, from the first 32 samples to the 224 samples after the starting point, were selected as the domain for feature extraction, based on the extracted inflection point.

For odor detection, the sensor response to the input odor was applied to the information entropy, as defined by C. Shannon, for the determination of the inflection point of the sensor measurement values on the input odor. The entropy in the frame unit for an individual sensor can be calculated from (1), where $p(i)$ is the probability of a certain range i, occurring in one frame. The occurrence rate $c(i)$ of the range i per frame is divided by the number of samples s per frame, as shown in (2), where i can be shown as $0 < i < L - 1$, implying that

the response range of a sensor is divided into multiples of L. Consider

$$E(n) = -\sum_{i=0}^{L-1} p(i) \log_2 p(i), \quad (1)$$

$$p(i) = \frac{c(i)}{s}. \quad (2)$$

Figure 5 shows the inflection point extraction procedure using entropy. Each description is as follows: (a) multiple odors pattern, (b) entropy value for each sensor signal, (c) number of sensors with entropy exceeding the threshold of 1.0, (d) response frame with more than seven sensors with entropies exceeding the 1.0 threshold, (e) result frame with the number of continuous frames of more than four frames, and (f) inflection point of multiple odors. In order to examine the changes in the state of odors, a hierarchical elimination method was employed to search for such changes.

The hierarchical elimination method determines how the present odor pattern has changed using the information on the odor pattern's starting and endpoint amplitudes along with information on the previous odor pattern conditions. When an odor is inputted, the amplitude increases, and when it is removed, the amplitude decreases. Thus, when the starting amplitude is smaller than that of the endpoint, an odor has been inputted, and when the starting point amplitude is larger than that of the endpoint, an odor has been removed. Multiple odors, a single odor, and an odorless state are determined according to the results of a conversion from the previous odor pattern to the present condition. A single odor is recognized when one odor is input into an odorless condition and when one of the two odors is removed. The conditions for multiple odors include that when multiple odors are being input into an odorless condition and that when another odor is input into a single odor. The condition needed to determine that an odorless state occurs when, during the initial odorless condition, one odor is removed from a single odor condition or when multiple odors are removed during the multiple odors condition. Thus, we searched for conditions such as the following: when an odorless state becomes a single odor or multiple odors, when a single odor is changed to an odorless state or to multiple odors, and when multiple odors are changed to an odorless state or to a single odor. The hierarchical elimination method decreases the number of recognition candidates.

Figure 6 shows the tree structure of multiple odors. When estimating the direction of odor sources, we did not consider the inflection point. This is the average amplitude of all sensors based on the signals of all 32 channels, which consists of a twin-sensor array of 16 channels. The sensor arrays were arranged so that a gap occurred between them, similar to the arrangement of the human nose. The amplitude of each sensor varies according to the direction of the odor. Figure 7 shows an example of the response of a real sensor based on this concept: (a) left sensor array and (b) right sensor array. We can confirm that the amplitude changes rapidly according to the near direction of the odor source.

Equations (3), (4), and (5) comprise the direction estimation method based on the amplitude of each channel.

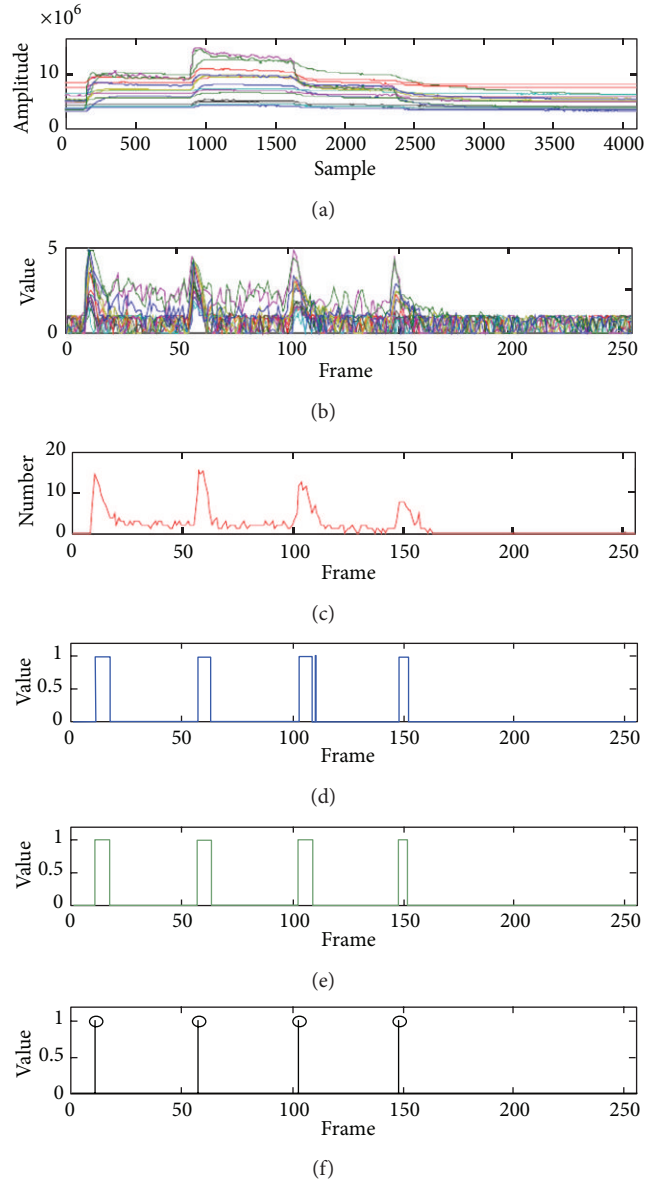


Figure 5: Inflection point extraction procedure using entropy.

D indicates a strength average of a sensor array, S indicates a sensor array, and n refers to the number of sensors

$$D = -\frac{1}{n}\sum_{n=1}^{N}\left(\frac{1}{T}\left(\sum_{t=t_0-T}^{t_0} S_{n_t}\right)\right), \quad (3)$$

$$O_{\text{source}L} \longleftarrow (D_L - D_R) > \varepsilon, \quad (4)$$

$$O_{\text{source}R} \longleftarrow (D_L - D_R) < -\varepsilon. \quad (5)$$

4. Experimental Results

We developed a 16-channel twin-sensor array. Table 1 shows the specifications of each sensor, including the Calefaction Voltage (V_H), Loop Voltage (V_C), Load Resistance (R_L), and Sensor Resistance (R_S). These factors were referenced by

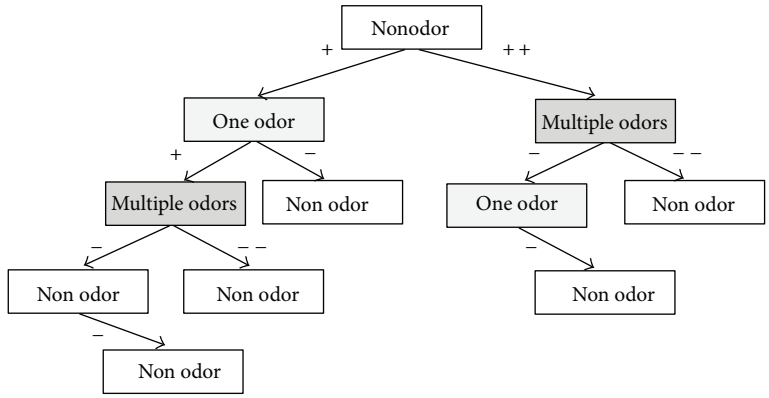
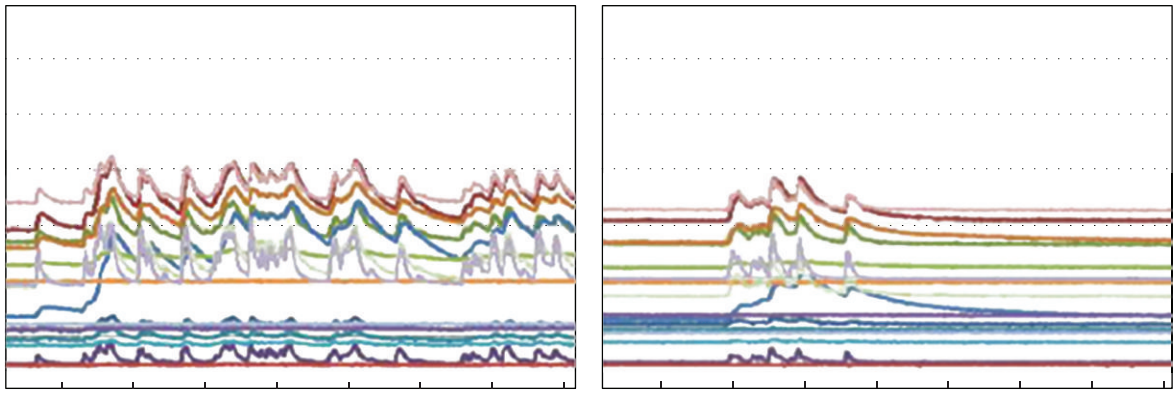


Figure 6: Tree structure of hierarchical elimination method.



(a) Left (b) Right

Figure 7: Changes of amplitude according to odor source direction.

the provider of each sensor. The structure of a sensor array can vary according to the target odor; thus, we experimented with artificial floral scents. Odor direction estimation is irrelevant to the target element of a sensor because it only uses the amplitude of an odor signal.

The objective of the experiments was to collect multiple odor databases against various cases in which multiple odors occurred and were then eliminated. We thus performed experiments focusing on detecting and recognizing odors using collected databases. We selected 11 scenarios containing various occurrences and eliminations of multiple odors, whereby multiple odor databases were collected. Due to the lack of internationally standardized odor databases, for our odor database collection we needed to use experimental environments, experimental conditions, experimental odor sources, and so forth. In this experiment (Figure 8), flower odor sources were selected that generated odors safely and evenly. As experimental flower odor sources, four types of flower odor oils were used, contained in a small 5 mL bottle. To evaluate the efficiency of the suggested system, a total of four types of floral fragrances were used.

The entire database was created by collecting 132 different combinations of scents, resulting from the double multiplication of six types of multiple odors by the 11 different scenarios. An example of the collection procedure for the odor database

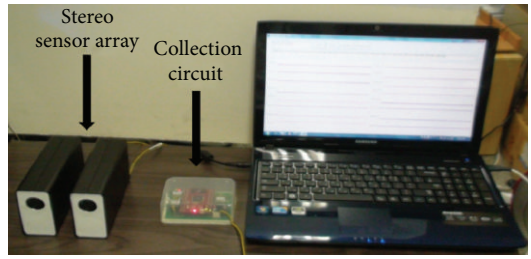


Figure 8: Components of proposed olfaction system for experiment.

collection is as follows: (1) start saving, (2) inputting odor ① at 10 sec, (3) inputting odor ② at 70 sec, (4) eliminating odor ② at 130 sec, (5) eliminating odor ① at 190 sec, (6) complete saving at 304 sec, and (7) ventilating. Odor detection and recognition experiments using multiple odor databases were performed. Odor detection was represented by the detection rate as an index of the performance evaluation of the number of times in which the inflection point was accurately detected. Odor recognition was represented by the recognition rate as an index of the performance evaluation of the number of times in which inputted odors were accurately recognized.

Table 1: Sensor specifications for the experiment.

No.	Sensor name	Target element	R_S (kilo Ω)	R_L (kilo Ω)	V_H (voltage)	V_C (voltage)
1	MiCS-5132	CO	16–160	24	3.3	2.5
2	MiCS-2710	NO_2	0.8–8	2	1.8	2.5
3	MiCS-5135	VOCs	20–400	100	3.3	2.5
4	MiCS-2601	O_3	3–60	10	2.5	2.5
5	MiCS-5521	CO, HC	24–1000	200	2.5	2.5
6	MQ-3	Alcohol	100–1000	200	5	2.5
7	GSAP61	VOCs	5–20	20	5	2.5
8	TGS-825	Hydrogen	3–30	12	5	2.5
9	MQ-6	LPG	10–60	20	5	2.5
10	GSET11	CO	70–811	200	5	2.5
11	MQ-7	CO	2–20	10	5	2.5
12	SP3-AQ2	VOCs	10–40	20	5	5
13	MiCS-5131	Ethanol	18–180	50	3.3	2.5
14	NGSX-03	Alcohol	10–50	20	5	2.5
15	MICS-4514 (1)	NO_2	0.8–8	2	5	2.5
16	MICS-4514 (2)	CO	100–1000	200	5	2.5

Table 2: Experimental results for odors recognition.

Details	Number of times	Success count	Recognition rate (%)
Mix ①②	18	16	88.89
Mix ①③	18	17	94.45
Mix ①④	18	18	100
Mix ②③	18	15	83.34
Mix ②④	18	15	83.34
Mix ③④	18	16	88.89
①	42	37	88.09
②	42	28	66.67
③	42	37	88.09
④	42	34	80.95
Odorless	132	130	98.48
Total	408	363	88.97

Table 3: Experimental results for estimation of direction.

Source direction	Sensor array	Initial response time	Estimation count	Estimation result
L	L R	20.45 30.05	38	$L:36$ $C:2$ $R:0$
C	L R	86.5 258.45	38	$L:7$ $C:27$ $R:4$
R	L R	23.25 90.5	38	$L:0$ $C:1$ $R:37$

Table 2 shows the odor recognition experiment results. The results of the odor recognition experiment showed a recognition rate of about 88.97%. However, the results of the odor recognition experiment that excluded the incorrectly detected results in the detection stage showed a recognition rate of about 92.84%. Based on the experimental results for odor recognition, the lower odor recognition rate is because the odor sources used in the experiment were all derived from flower oils. Although flower oils were used to ensure the safety of the experiment, they are vegetable based, with similar intensities and characteristics. However, when mixing two flower odors, the recognition rate was relatively high due to the changes in the intensity and characteristics of the odors. The recognition rate of the odorless state was not 100% due to misdetection. The recognition rate of the odorless state excluding misdetection was 100%.

Table 3 shows the experimental results for the estimation of the direction of the odor source. L, C, and R indicate the left, center, and right, respectively. The distance between the odor acquisition device and the floral scent was 1 m. In this estimation experiment, the odor sensor array was fixed. If this system was attached to a roll-around robot, it could search for and move to the location of the odor source. In the case of a center direction odor source, the recognition rate was somewhat poor, given the structure of the twin-sensor array. To improve the recognition rate for a center space, the distance between the sensor arrays should be adjusted according to the recognition space, or the number of sensor arrays should be increased.

5. Conclusion

In this study, we proposed a mechanical olfaction system composed of multiple odor recognition and odor source direction estimation abilities. The proposed system recognizes two odors simultaneously and estimates the direction of an odor source using signal amplitude. The proposed system is based on a 16-channel twin-sensor array, similar to the human olfactory system. We considered an entropy-based inflection point extraction method to detect odor activity, and proposed a hierarchical elimination method using PCA to recognize multiple odors. In addition, we used the amplitude of a stereo odor signal to estimate odor source direction. This study can be applied in various industries, study fields, robots, and realistic content in virtual reality.

The proposed olfaction system can be applied to various real-world environments. In the case of escaped gas, the proposed olfaction system can estimate the direction of a gas leak and the chemical ingredients. Moreover, the proposed system can be used as a method to distinguish between stale and fresh foods. In addition, the proposed system can detect drugs and chemicals and determine the concentration of the target matter. In future research, we intend to develop a roll-around humanoid robot with the proposed olfaction system capable of recognizing odors and of moving towards the various odor sources.

Acknowledgments

This research was supported by MSIP, Korea, under ITRC NIPA-2013-(H0301-13-3001) and Basic Science Research Program through NRF of Korea, funded by MOE (NRF-2010-0020210).

References

[1] W. Goepel, "Chemical imaging: I. Concepts and visions for electronic and bioelectronic noses," *Sensors and Actuators B*, vol. 52, no. 1-2, pp. 125–142, 1998.

[2] T. C. Pearce, S. S. Shciffman, H. T. Nagle, and J. W. Gardner, *Handbook of Machine Olfaction: Electronic Nose Technology*, Wiley-VCH, New York, NY, USA, 2003.

[3] H. K. Ekenel and B. Sankur, "Feature selection in the independent component subspace for face recognition," *Pattern Recognition Letters*, vol. 25, no. 12, pp. 1377–1388, 2004.

[4] Y.-W. Roh, D.-K. Kim, H.-O. Kwon, and K.-S. Hong, "A method of optimal sensor decision for odor recognition," *Sensor Letters*, vol. 9, no. 1, pp. 223–227, 2011.

[5] B.-G. Cheon, Y.-W. Roh, D.-J. Kim, and K.-S. Hong, "An implementation of floral scent recognition system using ica combined with correlation coefficients," in *Independent Component Analysis and Signal Separation* , vol. 5441 of *Lecture Notes in Computer Science*, pp. 654–661, 2009.

[6] S. K. Jha and R. D. S. Yadava, "Denoising by singular value decomposition and its application to electronic nose data processing," *IEEE Sensors Journal*, vol. 11, no. 1, pp. 35–44, 2011.

[7] C. Zanchettin, L. M. Almeida, F. D. Menezes, T. B. Ludermir, and W. M. Azevedo, "Odor recognition systems for natural gas odorization monitoring," in *Proceedings of the International Joint Conference on Neural Networks (IJCNN '12)*, pp. 1–8, 2012.

[8] P. Strobel, A. Lfakir, M. Siadat, and M. Lumbreras, "Detection of three pollutant gases and their mixtures in a humid atmosphere using a portable electronic nose," *Sensor Letters*, vol. 4, no. 3, pp. 222–228, 2006.

[9] A. Forleo, A. M. Taurino, S. Capone et al., "Design of an electronic nose for selective phosphine detection in cereals," *Sensor Letters*, vol. 4, no. 3, pp. 229–234, 2006.

[10] A. Lamagna, S. Reich, D. Rodríguez, and A. Boselli, "Hake freshness assessment with an electronic nose," *Sensor Letters*, vol. 4, no. 3, pp. 253–256, 2006.

[11] D.-K. Kim, Y.-W. Roh, and K.-S. Hong, "A method of multiple odors detection and recognition," in *Human-Computer Interaction. Interaction Techniques and Environments*, vol. 6762 of *Lecture Notes in Computer Science*, no. 2, pp. 464–473, 2011.

[12] V. V. Sysoev, E. Strelcov, and A. Kolmakov, "Multisensor microarrays based on metal oxide nanowires for electronic nose applications," in *Integrated Analytical Systems*, chapter 15, pp. 465–502, Springer, New York, NY, USA, 2013.

[13] A. Tripathy, A. K. Mohanty, and M. Narayan Mohanty, "Electronic nose for black tea quality evaluation using kernel based clustering approach," *International Journal of Image Processing*, vol. 6, no. 2, pp. 86–93, 2012.

[14] S. Vallone, N. W. Lloyd, S. E. Ebeler, and F. Zakharov, "Fruit volatile analysis using an electronic nose," *Journal of Visualized Experiments*, vol. 61, Article ID e3821, 2012.

[15] L. Zhang, F. Tian, L. Dang, and G. Li, "A novel ANN ensemble and self-calibration model in electonic nose for concentration estimation," in *Proceeding of the International Conference on Computer, Networks and Communication Engineering*, pp. 114–116, 2013.

[16] K. Brudzewski, S. Osowski, and A. Dwulit, "Recognition of coffee using differential electronic nose," *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 6, pp. 1803–1810, 2012.

[17] D. Wang, X. Wang, T. Liu, and Y. Liu, "Prediction of total viable counts on chilled pork using an electronic nose combined with support vector machine," *Meat Science*, vol. 90, no. 2, pp. 373–377, 2012.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 304631, 9 pages
http://dx.doi.org/10.1155/2013/304631

Research Article

Target Localization in Wireless Sensor Networks for Industrial Control with Selected Sensors

Zhenxing Luo,[1] Paul S. Min,[1] and Shu-Jun Liu[2]

[1] *Department of Electrical and Systems Engineering, Washington University at St. Louis, St. Louis, MO 63130, USA*
[2] *College of Communication Engineering, Chongqing University, Chongqing 400044, China*

Correspondence should be addressed to Zhenxing Luo; mariolzx@gmail.com

Received 19 January 2013; Accepted 28 May 2013

Academic Editor: Sabah Mohammed

This paper presents a novel energy-based target localization method in wireless sensor networks with selected sensors. In this method, sensors use Turbo Product Code (TPC) to transmit decisions to the fusion center. TPC can reduce bit error probability if communication channel errors exist. Moreover, in this method, thresholds for the energy-based target localization are designed using a heuristic method. This design method to find thresholds is suitable for uniformly distributed sensors and normally distributed targets. Furthermore, to save sensor energy, a sensor selection method is also presented. Simulation results showed that if sensors used TPC instead of Hamming code to transmit decisions to the fusion center, localization performance could be improved. Furthermore, the sensor selection method used can substantially reduce energy consumption for our target localization method. At the same time, this target localization method with selected sensors also provides satisfactory localization performance.

1. Introduction

With the advancement of hardware and software, industrial control systems (ICSs) have been widely used in many industrial areas. ICSs can use sensors to gather information about the environment or the condition of machines, and based on the information from sensors to monitor and control remote devices. Research efforts have focused on different areas of ICSs. For example, in [1], the authors discussed the security problem in ICSs. Research in [2] focused on the architecture and scheduling problem in ICSs while the authors in [3] presented an innovative control system for the industrial environment.

In many ICSs, a wireless sensor network (WSN) is laid out to control the robotics [4–6] or to track human motion [7]. Target position estimation is an essential part in these applications. There are many target localization methods available, such as direction of arrival (DOA) methods [8, 9], time-delay of arrival (TDOA) methods [10, 11], and energy-based target localization methods [12, 13]. The energy-based target localization method which uses quantized data in WSNs was presented in [12]. Compared with DOA methods and TDOA methods, energy-based target localization methods do not require sensors to have the ability to measure the direction of arrived signals, such as in DOA methods, or require accurate synchronization among sensors, such as in TDOA methods. Therefore, energy-based target localization methods are easier to implement than the DOA methods or TDOA methods [12]. This paper focuses on energy-based target localization methods.

In energy-based target localization methods, sensors measure the signals from the target and send the measurement information to the fusion center, which uses information from the sensors to estimate the target position [12]. Usually, sensors used in WSNs have limited sources, such as energy and communication bandwidth. Therefore, saving energy and communication bandwidth are very important in WSNs.

To save energy, sensors can quantize the information about the signals from the target before sending it to the fusion center [12, 14]. There are many quantization methods, such as the nonuniform quantization method in [15] and the vector quantization method in [16]. Interested readers can refer to [17]. In the quantization scheme, thresholds are the most important parameters. In [12], a heuristic method to determine the optimum quantization thresholds

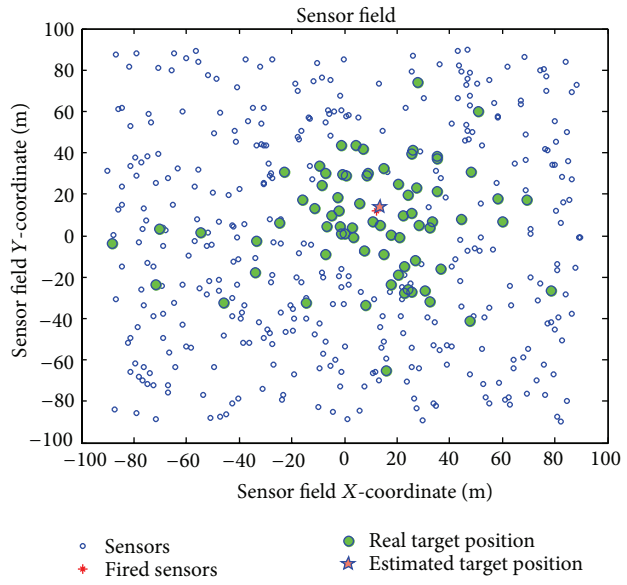


FIGURE 1: Sensor field (fired sensors are sensors having decisions other than 0).

was presented to calculate the optimum thresholds if sensor positions follow uniform distributions and the target position also follows uniform distributions. However, if the target is mostly present in a small area, Gaussian distribution can better describe the distributions of the target position. Our heuristic method can calculate optimum thresholds if the sensors are uniformly distributed and the target position follows Gaussian distributions (Figure 1).

In many applications, energy can also be saved by using a sensor selection scheme. In target tracking, [18] used an energy function and *posterior* Cramer-Rao lower bound (PCRLB) to select sensors. In [19], the authors used a modified Riccati equation approach to select sensors. In [20], a combination of PCRLB and sensor range was used to select sensors. In [21], the authors considered sensor selection when the detection probability of sensors was less than 1. More research in this area can be found in [22–25].

In target detection, [26] used a Kullback-Leibler (K-L) based approach to select sensors to maximize the K-L distance. In [27], an optimal sensor selection method in binary heterogeneous sensor networks was presented. In this paper, the K-L distance is maximized subject to cost constraints. In [28], the authors presented robust sensor selection methods to deal with uncertainties of the distribution mean.

In target estimation, if MLE is used, the Cramer-Rao lower bound (CRLB) or PCRLB can be used as the performance criterion. Then, sensors can be selected to maximize CRLB or PCRLB. In [29], sensors are selected based on PCRLB. However, in [29], several sensors were selected from all sensors. The computation cost of this selection method may be prohibitive if the total number of sensors is too large. To alleviate the computation cost, we propose a new sensor selection method in this paper.

The main contribution of this paper is a novel energy-based target localization method in WSNs with selected sensors. In this method, sensors use Turbo Product Code (TPC) to transmit decisions to the fusion center. Moreover, the thresholds used in our target localization method are determined using a new heuristic method specifically designed for sensor position following uniform distributions and target position following Gaussian distributions. Furthermore, a sensor selection method is presented, and this selection method can significantly save sensor energy without substantially sacrificing localization performance.

The rest of this paper is organized as follows. In Section 2, an energy-based target localization method using TPC is presented. A heuristic method to determine optimum thresholds is given in Section 3, and a novel sensor selection method is presented in Section 4. The analysis of performance loss and energy saving due to sensor selection is given in Section 5. The simulation setup is presented in Section 6, and the results and analysis are provided in Section 7. The concluding remarks are made in Section 8.

2. Energy-Based Target Localization Method Using TPC

The energy-based target localization method using quantized data was presented in [12]. Our energy-based target localization method is based on the method in [12]. However, in our method, TPC code and sensor selection are used. Also, we only estimate two parameters.

Following the derivation in [12], acoustic signal from a target decays as distance from the target to the measurement location increases. The relation can be determined by

$$a_i^2 = \frac{G_i P_0'}{(d_i/d_0)^n}. \tag{1}$$

In (1), P_0' is the power emitted by the target measured at a reference distance d_0, a_i is the signal amplitude from the target measured at the ith sensor, G_i is the gain of the ith sensor, which is determined by the sensor antenna, and n is the power decay exponent. The Euclidean distance between the target and the ith sensor is

$$d_i = \sqrt{(x_i - x_t)^2 + (y_i - y_t)^2}, \tag{2}$$

where (x_i, y_i) and (x_t, y_t) are the positions of sensor i and the target, respectively. It is assumed in this paper that every sensor has equal signal gain and $d_0 = 1$. The total number of sensors is denoted by N. Experiments showed that $n = 2$ [12]. Therefore, (1) can be simplified as

$$a_i^2 = \frac{P_0}{d_i^2}. \tag{3}$$

Because of the presence of noise in the environment, the signal that arrived at the ith sensor will be inevitably corrupted by noises. The process can be modeled and expressed as

$$s_i = a_i + w_i, \tag{4}$$

where w_i is a Gaussian noise following the distribution $N(0, \sigma^2)$. To save communication bandwidth and sensor

energy, s_i is quantized [12]. In the quantization process, the ith sensor quantizes the measured signal s_i into m_i according to a set of thresholds:

$$\overline{\gamma}_i = \left[\gamma_{i1}, \gamma_{i2}, \gamma_{i3}, \ldots, \gamma_{i(L-1)}, \gamma_{iL}\right]. \tag{5}$$

The quantization process used can be expressed as

$$m_i = \begin{cases} 0 & -\infty < s_i < \gamma_{i1} \\ 1 & \gamma_{i1} < s_i < \gamma_{i2} \\ 2 & \gamma_{i2} < s_i < \gamma_{i3} \\ \vdots & \vdots \\ L-2 & \gamma_{i(L-2)} < s_i < \gamma_{i(L-1)} \\ L-1 & \gamma_{i(L-1)} < s_i < \infty. \end{cases} \tag{6}$$

The probability that the transmitted decision m_i by the ith sensor takes value m can be calculated by

$$p\left(m_i = m \mid \theta\right) = \int_{(\eta_i - a_i)/\sigma}^{(\eta_{i+1} - a_i)/\sigma} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt. \tag{7}$$

The channel-aware target localization method was presented in [14]. However, the authors did not use any coding method to counter communication channel errors. In our method, to minimize the effect of communication channel errors, sensors use TPC to transmit decisions to the fusion center. Previous research has focused on using Hamming code to encode decisions from sensors to the fusion center for distributed detection [30–33]. In our method, sensors use TPC to transmit decisions to the fusion center. TPC is a category of Turbo codes and can provide an efficient way to construct long codes from short linear block ones with relatively high code rates [34–36].

In TPC, BCH code is used as the component codes and iterative soft-input soft-output algorithm is used to decode rows and columns code words. Then, we can determine the relation between $\widetilde{m}_i$ and m. If the relation between $\widetilde{m}_i$ and m can be determined, we can incorporate the transition relation into the MLE framework using an approach similar to that used in [14]. The system diagram is shown in Figure 2. After the transition probabilities from m to $\widetilde{m}_i$ are determined, the probability that $\widetilde{m}_i$ assumes the value m is given by

$$p(\widetilde{m}_i = m \mid \theta) = \sum_{m_i \in \{-1,1\}} p\left(\widetilde{m}_i = m \mid m_i\right) p(m_i \mid \theta). \tag{8}$$

In (8), $p(m_i \mid \theta)$ is defined in (7). If the decision vector received at the fusion center is $\widetilde{\mathbf{M}} = [\widetilde{m}_1\ \widetilde{m}_2 \cdots \widetilde{m}_N]^T$, the fusion center can find $\theta = [x_t \quad y_t]^T$ to maximize

$$\ln p\left(\widetilde{\mathbf{M}} \mid \theta\right) = \sum_{i=1}^{N} \ln \left[\sum_{m_i=0}^{L-1} p\left(\widetilde{m}_i = m \mid m_i\right) p\left(m_i \mid \theta\right)\right]. \tag{9}$$

The maximum likelihood estimator (MLE) tries to find $\widehat{\theta}$ to maximize

$$\widehat{\theta} = \max_{\theta} \ln p\left(\widetilde{\mathbf{M}} \mid \theta\right). \tag{10}$$

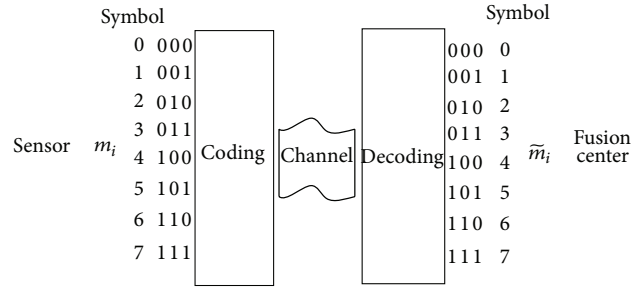


Figure 2: Symbol coding and decoding.

If an unbiased estimate $\widehat{\theta}$ exists, the CRLB of the estimate can be calculated by

$$E\left\{\left[\widehat{\theta}\left(\widetilde{\mathbf{M}}\right) - \theta\right]\left[\widehat{\theta}(\widetilde{\mathbf{M}}) - \theta\right]^T\right\} \geq \mathbf{J}^{-1}$$
$$\mathbf{J} = -E\left[\nabla_\theta \nabla_\theta^T \ln\ p\left(\widetilde{\mathbf{M}} \mid \theta\right)\right]. \tag{11}$$

The $\mathbf{J}$ matrix corresponding to our two parameters estimation problem can be denoted as a 2 by 2 matrix. The process to derive the $\mathbf{J}$ matrix is similar to the process in [14].

First, $\mathbf{J}(1,1)$ is derived:

$$\frac{\partial^2 \ln p\left(\widetilde{\mathbf{M}} \mid \theta\right)}{\partial x_t^2} = \sum_{i=1}^{N}\sum_{l=1}^{L-1} -\frac{\delta\left(\widetilde{m}_i - l\right)}{p^2\left(\widetilde{m}_i \mid \theta\right)}\left[\frac{\partial \ln p\left(\widetilde{m}_i \mid \theta\right)}{\partial x_t}\right]^2 + \frac{\delta\left(\widetilde{m}_i - l\right)}{p\left(\widetilde{m}_i \mid \theta\right)} \frac{\partial^2 p\left(\widetilde{m}_i \mid \theta\right)}{\partial x_t^2}. \tag{12}$$

Because the expectation of the second term of (12) is 0, the expectation of (12) is

$$E\left[\frac{\partial^2 \ln p\left(\widetilde{\mathbf{M}} \mid \theta\right)}{\partial x_t^2}\right] = \sum_{i=1}^{N}\sum_{l=1}^{L-1} -\frac{1}{p^2\left(\widetilde{m}_i \mid \theta\right)}\left[\frac{\partial p\left(\widetilde{m}_i \mid \theta\right)}{\partial x_t}\right]^2. \tag{13}$$

In (13), $p(\widetilde{m}_i \mid \theta)$ is defined in (8). The derivative of $p(\widetilde{m}_i \mid \theta)$ with respect to x_t is

$$\frac{\partial p\left(\widetilde{m}_i \mid \theta\right)}{\partial x_t} = \sum_{m_i=0}^{L-1} p\left(\widetilde{m}_i \mid m_i\right) \frac{\partial p\left(m_i \mid \theta\right)}{\partial x_t}. \tag{14}$$

In (14), $\partial p(m_i \mid \theta)/\partial x_t$ can be obtained as

$$\begin{aligned}\frac{\partial p\left(m_i = l \mid \theta\right)}{\partial x_t} &= \frac{\partial}{\partial x_t} \\ &\quad \times \left[Q\left(\frac{\eta_{il} - a_i}{\sigma}\right) - Q\left(\frac{\eta_{i(l+1)} - a_i}{\sigma}\right)\right] \\ &= \frac{-\sqrt{P_0}\left(x_t - x_i\right)}{d_i^3\sqrt{2\pi}\sigma} \\ &\quad \times \left(e^{-(\eta_{il}-a_i)^2/2\sigma^2} - e^{-(\eta_{i(l+1)}-a_i)^2/2\sigma^2}\right).\end{aligned} \tag{15}$$

Then, (14) can be expressed as

$$\begin{aligned}&\frac{\partial p\left(\widetilde{m}_i \mid \theta\right)}{\partial x_t} \\ &\quad = \sum_{m_i=0}^{L-1} p\left(\widetilde{m}_i \mid m_i\right) \frac{-\sqrt{P_0}\left(x_t - x_i\right)}{d_i^3\sqrt{2\pi}\sigma} \\ &\qquad \times \left(e^{-(\eta_{il}-a_i)^2/2\sigma^2} - e^{-(\eta_{i(l+1)}-a_i)^2/2\sigma^2}\right).\end{aligned} \tag{16}$$

Other elements of **J** can be calculated similarly.

3. A Heuristic Quantization Method to Determine Optimum Threshold

Before the MLE method can be used, we have to determine an appropriate set of thresholds. Thresholds cannot be too high; otherwise, sensors will make the same decision, 0, and the MLE method cannot proceed. On the other hand, thresholds cannot be too low; otherwise, all sensors will report decision $L - 1$ and the MLE method also cannot be implemented. In our experiments, sensors are uniformly distributed across the field. The X-coordinate of the target follows Gaussian distribution $N(0, \sigma_x^2)$ and the Y-coordinate of the target follows Gaussian distribution $N(0, \sigma_y^2)$. Now the detailed steps to determine the optimum threshold are presented. The method is similar to the method used in [12].

First, we assume that the sensor position x_i follows uniform distribution

$$x_i \sim u\left[-a, a\right], \tag{17}$$

and the probability density function (PDF) of x_i is

$$f\left(x_i\right) = \frac{1}{2a}. \tag{18}$$

Similarly, we assume that the sensor position y_i also follows uniform distribution

$$y_i \in u\left[-a, a\right], \tag{19}$$

and the PDF of y_i is

$$f\left(y_i\right) = \frac{1}{2a}. \tag{20}$$

Then, we know $-x_i$ also follows uniform distribution $-x_i \sim u[-a, a]$ and the PDF of $-x_i$ is

$$f\left(x_i\right) = \frac{1}{2a}. \tag{21}$$

Moreover, the target position x_t follows Gaussian distribution $x_t \sim N(0, \sigma_x^2)$, and the PDF of x_t is

$$f\left(x_t\right) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-x_t^2/2\sigma_x^2}. \tag{22}$$

Similarly, the target position y_t follows Gaussian distribution $y_t \sim N(0, \sigma_y^2)$ and the PDF of y_t is

$$f\left(y_t\right) = \frac{1}{\sqrt{2\pi}\sigma_y} e^{-x_t^2/2\sigma_y^2}. \tag{23}$$

We denote t as $t = x_i - x_t$. Because the PDF of the sum of two independent variables is the convolution of their PDFs, we can have

$$f_T\left(t\right) = \int_{t-a}^{t+a} \frac{1}{\sqrt{2\pi}\sigma_y} e^{-\tau^2/2\sigma_y^2} \frac{1}{2a} d\tau. \tag{24}$$

Then, (24) can be expressed as

$$f_T\left(t\right) = \frac{1}{2a}\left(Q\left(t - a\right) - Q\left(t + a\right)\right), \quad t \in (-\infty, \infty), \tag{25}$$

in which $Q(x)$ is defined as

$$Q\left(x\right) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt. \tag{26}$$

If we define $u = (x_i - x_t)^2$, we have $u = t^2$, and $f(u)$ is

$$f\left(u\right) = \frac{1}{2\sqrt{u}} \frac{1}{2a}\left(Q\left(\sqrt{u} - a\right) - Q\left(\sqrt{u} + a\right)\right), \quad u \in (0, \infty). \tag{27}$$

If we define $v = (x_i - x_t)^2 + (y_i - y_t)^2$, we can have

$$\begin{aligned}&f\left(v\right) \\ &\quad = \int_0^v \frac{1}{4a\sqrt{\tau}}\left(Q\left(\sqrt{\tau} - a\right) - Q\left(\sqrt{\tau} + a\right)\right) \\ &\qquad \times \frac{1}{4a\sqrt{v - \tau}}\left(Q\left(\sqrt{v - \tau} - a\right) - Q\left(\sqrt{v - \tau} + a\right)\right) d\tau.\end{aligned} \tag{28}$$

Because sensors have to be at least d_0 meters away from the target, the minimum value of v is d_0^2. The probability that v is greater than d_0^2 is

$$\alpha = 1 - \int_0^{d_0^2} f\left(v\right) dv. \tag{29}$$

Then, if v is greater or equal to d_0^2, we have

$$f_v\left(v \mid v \geq d_0^2\right) = \frac{1}{\alpha} f_v\left(v\right). \tag{30}$$

If we denote w as $w = P_0/v$, we can have

$$f(w) = \frac{P_0}{w^2} f_v\left(\frac{P_0}{w}\right), \quad w \in \left(0, \frac{P_0}{d_0^2}\right). \tag{31}$$

Then, (31) can be expressed as

$$\begin{aligned} f(w) = {} & \frac{P_0}{w^2}\frac{1}{\alpha} \\ & \times \left[\int_0^{P_0/w} \frac{1}{4a\sqrt{\tau}} \frac{1}{4a\sqrt{(P_0/w) - \tau}} \right. \\ & \quad \times \left(Q\left(\sqrt{\tau} - a\right) - Q\left(\sqrt{\tau} + a\right)\right) \\ & \quad \times \left(Q\left(\sqrt{\frac{P_0}{w} - \tau} - a\right)\right. \\ & \quad \left.\left. -Q\left(\sqrt{\frac{P_0}{w} - \tau} + a\right)\right) d\tau \right]. \end{aligned} \tag{32}$$

If we define $z = \sqrt{w}$, $f(z)$ can be expressed as

$$f(z) = 2z f_w\left(z^2\right), \quad z \in \left(0, \sqrt{\frac{P_0}{d_0}}\right) \tag{33}$$

and finally as

$$\begin{aligned} f(z) = {} & 2z\frac{P_0}{z^4}\frac{1}{\alpha} \\ & \times \left[\int_0^{P_0/w} \frac{1}{4a\sqrt{\tau}} \frac{1}{4a\sqrt{(P_0/z^2) - \tau}} \right. \\ & \quad \times \left(Q\left(\sqrt{\tau} - a\right) - Q\left(\sqrt{\tau} + a\right)\right) \\ & \quad \times \left(Q\left(\sqrt{\frac{P_0}{z^2} - \tau} - a\right)\right. \\ & \quad \left.\left. -Q\left(\sqrt{\frac{P_0}{z^2} - \tau} + a\right)\right) d\tau \right]. \end{aligned} \tag{34}$$

Now, we have the expression $f(z)$. According to [8], optimum threshold can be chosen so that sensors will make a decision from 0 to $L-1$ with equal probability. The probability that a sensor will make decision l is

$$P_l = \int_0^{\sqrt{P_0}/d_0} \left[Q\left(\frac{\eta_l - z}{\sigma}\right) - Q\left(\frac{\eta_{(l+1)} - z}{\sigma}\right)\right] f(z)\, dz. \tag{35}$$

Using (35), the optimum set of thresholds can be determined.

4. Sensor Selection Method

Energy is a precious resource in WSNs. Battery-powered sensors may be deployed in a remote area and replacing batteries is infeasible in many situations. Therefore, saving energy is very important in WSNs. The energy a sensor uses can be divided into three parts. The first part E_1 is the energy a sensor uses to measure the signal from the target. The second part E_2 is the energy a sensor uses to maintain essential functions, such as receiving information from the fusion center and keeping itself awake. The third part E_3 is the energy a sensor uses to send the decisions to the fusion center. Because energy consumed in communication consists of the majority part of energy consumption, reducing E_3 can significantly save sensor energy and greatly extend the operation time of a sensor. A sensor selection scheme can reduce energy consumption by choosing sensors containing more useful information and allowing those sensors to send the decisions to the fusion center while sensors containing less useful information are not allowed to send decisions to the fusion center.

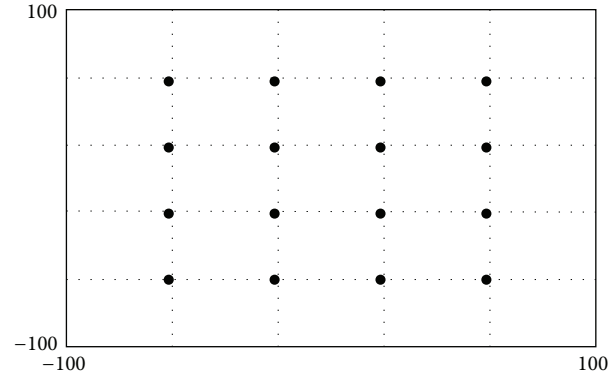


Figure 3: Sensor field divided into 25 sections (black points denote anchor sensors).

In the target estimation method in [29], sensors are selected based on PCRLB from all sensors. The computation cost of this selection method may be prohibitive if the total number of sensors is large. We propose a new sensor selection method, which can alleviate the computation cost.

The steps of our method are as follows.

(1) Divide the whole sensor field into different regions. Place N_0 number of anchor sensors. For example, anchor sensors can be placed into a grid as shown in Figure 3.

(2) Use anchor sensors and the weighted average method to estimate a coarse target position.

(3) Use the coarse target position to choose all sensors in the region where the estimated target is located. For example, in Figure 4, region 1 will be chosen.

(4) If the target falls into region M_1, sensors in the neighboring region, region 2, will also be chosen. If the target falls into the region M_2, sensors in the neighboring region, region 3, will also be chosen. If the target falls into M_3, sensors in all neighboring regions, region 2, region 3, and region 4, will also be chosen.

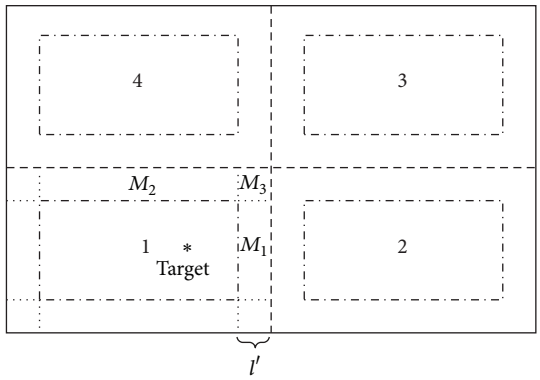


Figure 4: Example of sensor regions.

(5) Chosen sensors will report decisions to the fusion center. Sensors not in chosen regions will not report decisions.

Our sensor selection method is easy to implement and circumvents the computation problem in [29]. In our method, the whole sensor field is divided into many smaller regions and a coarse target position is used to select regions in which sensors need to report their decisions to the fusion center. Therefore, our method can greatly save sensor energy.

5. Performance Loss and Energy Saving

Although our method can save sensor energy, the target localization performance may suffer because the fusion center does not have data from all sensors. Therefore, we also want to know the performance loss and the amount of energy saved due to sensor selection. Knowing the performance loss and how much energy was saved can help us to make the decision about when to use the sensor selection method.

5.1. Performance Loss. In our method, only two parameters are estimated. Therefore, $\mathbf{J}$ is a 2 by 2 matrix:

$$\mathbf{J} = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix}. \tag{36}$$

If the set of selected sensors is R_1 and the set of nonselected sensors is R_2, then every element of $\mathbf{J}$ can be divided into two parts. For example,

$$\begin{aligned} \mathbf{J}(1,1) &= \sum_{i=1}^{N}\sum_{l=1}^{L-1} - \frac{1}{p^2(\widetilde{m}_i \mid \theta)} \left[\frac{\partial p(\widetilde{m}_i \mid \theta)}{\partial x_t}\right]^2 \\ &= \sum_{i\in R_1}\sum_{l=1}^{L-1} - \frac{1}{p^2(\widetilde{m}_i \mid \theta)} \left[\frac{\partial p(\widetilde{m}_i \mid \theta)}{\partial x_t}\right]^2 \\ &\quad + \sum_{i\in R_2}\sum_{l=1}^{L-1} - \frac{1}{p^2(\widetilde{m}_i \mid \theta)} \left[\frac{\partial p(\widetilde{m}_i \mid \theta)}{\partial x_t}\right]^2 \\ &= \mathbf{J}^1(1,1) + \mathbf{J}^2(1,1) \\ \mathbf{J} &= \mathbf{J}^1 + \mathbf{J}^2 = \begin{bmatrix} J_{11}^1 & J_{12}^1 \\ J_{21}^1 & J_{22}^1 \end{bmatrix} + \begin{bmatrix} J_{11}^2 & J_{12}^2 \\ J_{21}^2 & J_{22}^2 \end{bmatrix}. \end{aligned} \tag{37}$$

Similarly, other elements can also be expressed as two parts: fisher information contributed by selected sensors and fisher information contributed by nonselected sensors. The performance loss is due to loss of Fisher information from nonselected sensors. However, this method to determine performance loss is valid only if the target is at a fixed position. If the target is randomly distributed, root-mean-square (RMS) estimation errors can be used to show the performance loss. We can compare the RMS estimation errors provided by the MLE method using data from all sensors and the RMS estimation errors provided by the MLE method using data only from chosen sensors.

5.2. Energy Saving. It is assumed that each region has the same size and the same M_1, M_2, and M_3. We use A to denote the area of each region, A_1 to denote the area of M_1, A_2 to denote the area of M_2, and A_3 to denote the area of M_3. Also, we assume that the fusion center is far away from the sensors. Therefore, the difference in sensor positions is not important and each sensor will consume the same amount of energy to send decisions to the fusion center. This energy is denoted by e_1, which can be calculated by

$$e_1\left(m, d_{f,i}\right) = E_{\text{elec}} \times m + \varepsilon_{\text{amp}} \times m \times d_{f,i}^2, \tag{38}$$

where $E_{\text{elec}} = 50$ uJ/bit, $\varepsilon_{\text{amp}} = 100$ nJ/bit/m^2, and $d_{f,i}$ is the distance between the ith sensor and the fusion center [37].

The total number of sensors in the sensor field is N, and we assume that each region has the same number of sensors, N_{region}. Then, the probability that the target is located in region M_1 is A_1/A, the probability that the target is located in M_2 is A_2/A, and the probability that the target is located in M_3 is A_3/A. We assume that the target mostly concentrates in the center of the field, and in the center of the field, the target is almost uniformly distributed. In each region, there are two M_1 areas, two M_2 areas, and four M_3 areas. If the target is located in M_1, the neighboring region is also activated. Similarly, If the target is located in M_2, the neighboring region is also activated. If the target is located in M_3, the other three neighboring regions are also activated. Therefore, the energy sensors consume to transmit decisions can be expressed as

$$\begin{aligned} E_4 &= e_1 * N_{\text{region}} \\ &\quad * \left(\frac{A - 2A_1 - 2A_2 - 4A_3}{A} + 4\frac{A_1}{A} + 4\frac{A_2}{A} + 16\frac{A_3}{A}\right). \end{aligned} \tag{39}$$

Without sensor selection, the energy consumed by sensors to transmit decisions is

$$E_5 = N_{\text{total}} * N_{\text{field}} * e_1. \tag{40}$$

Now, we know the performance loss and energy saved. Using the information about the performance loss and energy saved, we can strike the balance between performance loss and energy saving. Also, a multi-objective optimization method can be used to jointly optimize performance loss and energy saving.

6. Simulation Setup

6.1. TPC Encoding and Decoding. If sensors use TPC to transmit decisions to the fusion center, localization performance can be improved due to lower bit error probability. For comparison purposes, simulations were run to show the localization performance when TPC code was used and the localization performance when Hamming code was used. In simulations, each sensor quantized the received signal into 3 bits data. In hamming code, we used (7, 4, 3) Hamming code. In TPC, we used BCH (32, 26) codes as the component code, so the rate of TPC was $(26/32) \times (26/32) = 0.67$. A highly efficient iterative decoding algorithm, ALD-FBBA, is used at the receiver side. ALD-FBBA can provide good performance with low complexity. If communication channel SNR was 7 dB, the bit error probability was 0.0077 for Hamming code and 0.0002 for TPC code. We calculated the decision transition matrix using the bit error probability. Using the decision transition matrix corresponding to Hamming code and the decision transition matrix corresponding to the TPC, we compared the effect of coding on localization performance. We set $(x_t, y_t) = (0, 0)$, $a = 90$, and $P_0 = 8000$. We used 480 sensors and varied σ to see the effect of signal noise on localization performance. 100 Monte Carlo simulations were used to calculate RMS errors. The thresholds were set to

$$\overline{\gamma}_i = [-\infty, 0.917, 1.017, 1.116, 1.25, 1.45, 1.79, 2.6, \infty]. \tag{41}$$

We define RMS location error as

$$\Delta = \text{RMS errors}\left(\widehat{x}_t - x_t\right) + \text{RMS errors}\left(\widehat{y}_t - y_t\right) \tag{42}$$

and used Δ as performance criterion.

6.2. Performance Loss and Energy Saving. We used random target locations in the simulation. Therefore, we used Δ in (42) as the performance criterion. Energy saving can be easily calculated by comparing E_4 in (39) and E_5 in (40). In this simulation, sensors were uniformly distributed in the sensor field with $a = 90$ and the target position followed Gaussian distribution with $x_t \sim N(0, 36^2)$ and $y_t \sim N(0, 36^2)$. We set $\sigma = 1$, $P_0 = 8000$, and assumed perfect communication channels. We divided the sensor field into 25 regions as shown in Figure 3 and set the l' in Figures 2 to 4. The whole sensor field had 480 sensors. Each sensor used the same amount of energy, $E_1 + E_2 = 6e - 3$ J, every second whether it was selected or not. Selected sensors made 10,000 sets of transmission every second. We assumed that the distances from the sensors to the fusion center were the same, 1,000 meters. The optimum thresholds for this setting were

$$\overline{\gamma}_i = [-\infty, 0, 0.54, 0.96, 1.35, 1.76, 2.27, 3.31, \infty]. \tag{43}$$

The optimum thresholds were used in the simulations.

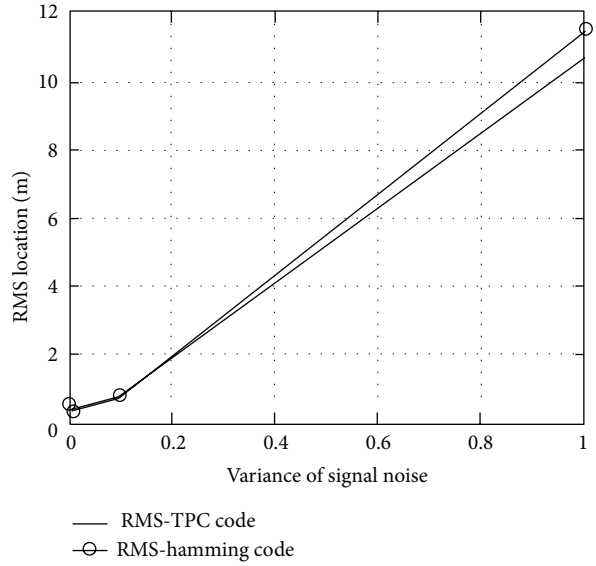


Figure 5: RMS location errors as a function of the variance of signal noise (RMS location errors were defined in (42), and the variance of signal noise is σ^2).

7. Results and Analysis

7.1. TPC Encoding and Decoding. Figure 5 shows that RMS location error as a function of the variance of signal noise. The RMS location error increased as the variance of signal noise increased. Moreover, the RMS location errors given by the MLE method using TPC were lower than the RMS location errors given by Hamming code. This means that if the MLE method uses TPC, the localization performance will be better. The threshold set used in this simulation was presented in (41), which was an optimum threshold set determined by the heuristic method. To determine the optimum set of thresholds, we used $a = 90$ and $\sigma_x^2 = \sigma_y^2 = 0.001^2$. The variance of x_t and the variance of y_t were very small. That means the target rarely moved to another position and was almost fixed at position $(0, 0)$. Therefore, the threshold set determined using these parameters was valid for our simulation where $(x_t, y_t) = (0, 0)$. Moreover, when the sensors used the optimum threshold set in (41), the numbers of sensors making each particular decision were almost the same (Figure 6). The results in Figure 6 validated the optimum set of thresholds calculated using (35).

7.2. Performance Loss and Energy Saving. Simulation results showed that using the setting described in Section 6.2, we had $\Delta_{\text{with\_selection}} = 8.3031$ and $\Delta_{\text{without\_selection}} = 3.9825$. The difference between $\Delta_{\text{with\_selection}}$ and $\Delta_{\text{without\_selection}}$ is performance loss due to sensor selection. As for energy saving, if our sensor selection method was used, the energy consumption of all sensors was $E_{\text{selection}} = 28.4928$ J per second. If our sensor selection was not used, the energy consumption of all sensors was $E_{\text{without\_selection}} = 483.12$ J per second. We can see that the sensor selection method

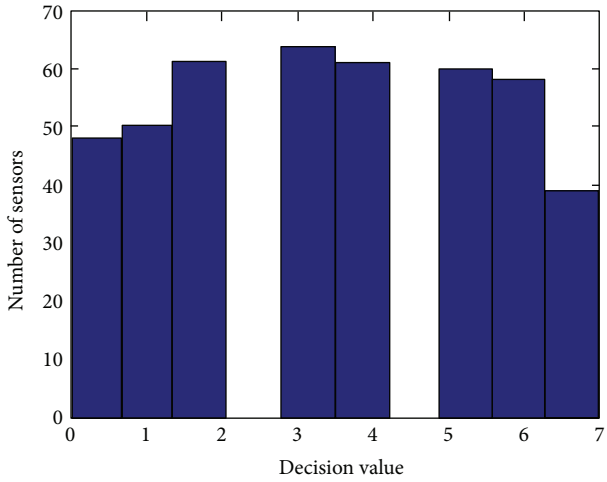


Figure 6: Number of sensors as a function of decision value.

can greatly reduce energy consumption while sacrificing a moderate degree of performance.

8. Conclusion

In this paper, a novel energy-based target localization method with selected sensors was presented. It is well known that communication channel errors can degrade the energy-based target localization methods. Our new target localization method uses TPC to counter communication channel errors. Simulation results showed that the target localization method using TPC can provide better localization performance than the target localization method using hamming code. Moreover, the sensor selection method used can save energy without great performance degradation as demonstrated by simulation results. Furthermore, if sensors use the optimum thresholds determined by the heuristic method, there will be about the same number of sensors making each particular decision. Overall, the new energy-based target localization method with selected sensors can achieve good localization performance with less energy consumption.

Acknowledgment

The authors want to thank Bingchao Liu from the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, China, for providing numerical parameters about Hamming code and TPC code.

References

[1] P. Jie and L. Li, "Industrial control system security," in *Proceedings of the 3rd International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC '11)*, pp. 156–158, Zhejiang, China, August 2011.

[2] J. Wang, H. Wang, H. Yu, and A. Xu, "Research of architecture and scheduling for wireless industrial control system," in *Proceedings of the IEEE International Conference on Information Acquisition (ICIA '06)*, pp. 936–940, Shandong, China, August 2006.

[3] M. Foursa, D. D'Angelo, and R. Narayanan, "Innovative control system for industrial environment," in *Proceedings of the 4th International Conference on Autonomic and Autonomous Systems (ICAS '08)*, pp. 82–87, March 2008.

[4] B. Song, G. Tian, G. Li, F. Zhou, and D. Liu, "ZigBee based wireless sensor networks for service robot intelligent space," in *Proceedings of the International Conference on Information Science and Technology (ICIST '11)*, pp. 834–838, Nanjing, China, March 2011.

[5] G. Mester, "Wireless sensor-based control of mobile robots motion," in *Proceedings of the 7th International Symposium on Intelligent Systems and Informatics (SISY '09)*, pp. 81–84, Subotica, Serbia, September 2009.

[6] G. Mester, I. Matijevics, T. Szepe, and J. Simon, *Application and Multidisciplinary Aspects of Wireless Sensor Networks: Concepts, Integration, and Case Studies*, Springer, New York, NY, USA, 2010.

[7] S. Zhang, W. Xiao, J. Gong, and Y. Yin, "A novel human motion tracking based on wireless sensor network," *International Journal of Distributed Sensor Networks*, vol. 2013, Article ID 636052, 7 pages, 2013.

[8] X. Li and J. Huang, "Bayesian high resolution DOA estimator based on importance sampling," in *Proceedings of the Oceans 2005—Europe*, pp. 611–615, June 2005.

[9] R. Sanudin, N. H. Noordin, A. O. El-Rayis, N. Haridas, A. T. Erdogan, and T. Arslan, "Analysis of DOA estimation for directional and isotropic antenna arrays," in *Proceedings of the 7th Loughborough Antennas and Propagation Conference (LAPC '11)*, pp. 1–4, Loughborough, UK, November 2011.

[10] M. Sun and K. C. Ho, "An asymptotically efficient estimator for TDOA and FDOA positioning of multiple disjoint sources in the presence of sensor location uncertainties," *IEEE Transactions on Signal Processing*, vol. 59, no. 7, pp. 3434–3440, 2011.

[11] K. C. Ho, "Bias reduction for an explicit solution of source localization using TDOA," *IEEE Transactions on Signal Processing*, vol. 60, no. 5, pp. 2101–2114, 2012.

[12] R. X. Niu and P. K. Varshney, "Target location estimation in sensor networks with quantized data," *IEEE Transactions on Signal Processing*, vol. 54, no. 12, pp. 4519–4528, 2006.

[13] X. Sheng and Y.-H. Hu, "Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 53, no. 1, pp. 44–53, 2005.

[14] O. Ozdemir, R. Niu, and P. K. Varshney, "Channel aware target localization with quantized data in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 57, no. 3, pp. 1190–1202, 2009.

[15] M. Hagiwara, T. Kitayabu, H. Ishikawa, and H. Shirai, "Delta-sigma modulator with non-uniform quantization," in *Proceedings of the IEEE Radio and Wireless Symposium (RWS '11)*, pp. 351–354, Phoenix, AZ, USA, January 2011.

[16] T. K. Grönfors and N. S. Päivinen, "Comparison of vector quantization methods for medical fidelity preserving lossy compression of EMG signals," in *Proceedings of the International Conference on Computational Intelligence for Modelling, Control and Automation (CIMCA '05) and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (IAWTIC '05)*, pp. 1107–1113, Vienna, Austria, November 2005.

[17] B. Widrow and I. Kollár, *Quantization Noise: Roundoff Error in Digital Computation, Signal Processing, Control, and Communications*, Cambridge University Press, Cambridge, UK, 2008.

[18] H. Ma, "Sensing nodes selection scheme for distributive target tracking in wireless sensor networks," in *Proceedings of the 37th Annual Conference of the IEEE Industrial Electronics Society (IECON '11)*, pp. 2164–2169, Melbourne, Australia, November 2011.

[19] U. Ramdaras and F. Absil, "Sensor selection: the modified Riccati equation approach compared with other selection schemes," in *Proceedings of the 10th International Conference on Information Fusion (FUSION '07)*, Quebec, Canada, July 2007.

[20] Z. Liu, J. Wang, and H. Wang, "PCRLB-based cluster selection for target tracking in wireless sensor networks," in *Proceedings of the 3rd International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC '11)*, pp. 83–87, Beijing, China, October 2011.

[21] U. D. Ramdaras and F. Bolderheij, "Performance-based sensor selection for optimal target tracking," in *Proceedings of the 12th International Conference on Information Fusion (FUSION '09)*, pp. 1687–1694, July 2009.

[22] F. R. Armaghani, I. Gondal, and J. Kamruzzaman, "Dynamic sensor selection for target tracking in wireless sensor networks," in *Proceedings of the IEEE 74th Vehicular Technology Conference (VTC Fall '11)*, San Francisco, CA, USA, September 2011.

[23] Z. Liu and J. Wang, "PCRLB-based optimal sensor selection for maneuvering target tracking," in *Proceedings of the 7th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM '11)*, Wuhan, China, September 2011.

[24] S. Zhang, W. Xiao, M. H. Ang Jr., and C. K. Tham, "IMM filter based sensor scheduling for maneuvering target tracking in wireless sensor networks," in *Proceedings of the International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP 07)*, pp. 287–292, Melbourne, Australia, December 2007.

[25] M. R. Zoghi and M. H. Kahaei, "Sensor selection for target tracking in WSN using modified INS algorithm," in *Proceedings of the 3rd International Conference on Information and Communication Technologies: From Theory to Applications (ICTTA '08)*, Damascus, Syria, April 2008.

[26] V. Srivastava, K. Plarre, and F. Bullo, "Adaptive sensor selection in sequential hypothesis testing," in *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC '11)*, pp. 6284–6289, Orlando, FL, USA, December 2011.

[27] M. Lázaro, M. Sánchez-Fernández, and A. Artes-Rodriguez, "Optimal sensor selection in binary heterogeneous sensor networks," *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1577–1587, 2009.

[28] D. Bajović, B. Sinopoli, and J. Xavier, "Sensor selection for event detection in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4938–4953, 2011.

[29] E. Maşazade, R. X. Niu, P. K. Varshney, and M. Keskinoz, "Energy aware iterative source localization for wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 58, no. 9, pp. 4824–4835, 2010.

[30] T.-Y. Wang, Y. S. Han, B. Chen, and P. K. Varshney, "A combined decision fusion and channel coding scheme for distributed fault-tolerant classification in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 5, no. 7, pp. 1695–1705, 2006.

[31] C. Yao, P.-N. Chen, T.-Y. Wang, Y. S. Han, and P. K. Varshney, "Performance analysis and code design for minimum hamming distance fusion in wireless sensor networks," *IEEE Transactions on Information Theory*, vol. 53, no. 5, pp. 1716–1734, 2007.

[32] T.-Y. Wang, Y. S. Han, and P. K. Varshney, "Fault-tolerant distributed classification based on non-binary codes in wireless sensor networks," *IEEE Communications Letters*, vol. 9, no. 9, pp. 808–810, 2005.

[33] T.-Y. Wang, Y. S. Han, P. K. Varshney, and P.-N. Chen, "Distributed fault-tolerant classification in wireless sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 4, pp. 724–734, 2005.

[34] J. Fang, F. Buda, and E. Lemois, "Turbo product code: a well suitable solution to wireless packet transmission for very low error rates," in *Proceedings of the 2nd International Symposium on Turbo Codes and Related Topics*, pp. 101–111, 2000.

[35] R. M. Pyndiah, "Near-optimum decoding of product codes: block turbo codes," *IEEE Transactions on Communications*, vol. 46, no. 8, pp. 1003–1010, 1998.

[36] M. P. C. Fossorier and S. Lin, "Soft-input soft-output decoding of linear block codes based on ordered statistics," in *Proceedings of the of IEEE Global Telecommunications Conference (GLOBECOM '98)*, pp. 2828–2833, November 1998.

[37] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proceedings of the 33rd Annual Hawaii International Conference on System Siences*, pp. 10–20, January 2000.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 471917, 14 pages
http://dx.doi.org/10.1155/2013/471917

Research Article

On the Impact of Local Processing for Motor Monitoring Systems in Industrial Environments Using Wireless Sensor Networks

Ruan Delgado Gomes,[1] Marcéu Oliveira Adissi,[2] Abel Cavalcante Lima-Filho,[2] Marco Aurélio Spohn,[3] and Francisco Antônio Belo[2]

[1] *The Federal Institute of Education, Science, and Technology of Paraíba, 58200-000 Guarabira, PB, Brazil*
[2] *The Federal University of Paraíba, 58051-900 João Pessoa, PB, Brazil*
[3] *The Federal University of Fronteira Sul, 89812-000 Chapecó, SC, Brazil*

Correspondence should be addressed to Ruan Delgado Gomes; ruandgomes@gmail.com

Received 1 February 2013; Accepted 5 June 2013

Academic Editor: Tai-hoon Kim

This paper presents a theoretical study for verifying the impact of using smart nodes in motor monitoring systems in industrial environments employing Wireless Sensor Networks (WSNs). Structured cabling and sensor deployment are usually more expensive than the cost of the sensors themselves. Besides the high cost, the wired approach offers little flexibility, making the network deployment and maintenance a complex process. In this context, wireless networks present a number of advantages compared to wired networks as, for example, the ease and speed of deployment and maintenance and the associated low cost. However, WSNs have several limitations, such as the low bandwidth and unreliability, especially in harsh environments (e.g., industrial plants). This paper presents a theoretical study on the performance of WSNs for motor monitoring applications in industrial environments, taking into account WSNs' characteristics (i.e., unreliability and communication and processing latency). The results obtained through mathematical models were analyzed together with experimental results, and it was demonstrated that employing intelligent nodes with local processing capabilities is essential for the applications under consideration, because it reduces the amount of data transmitted over the network allowing monitoring even in scenarios with high interference rate, paying off the extra latency resulting from local processing.

1. Introduction

The optimization of power usage, by decreasing costs and reducing environmental impact, has been of great concern to various sectors in the society. In this context, intelligent and low cost industrial automation systems for motor monitoring can be a very useful tool, by reducing electricity consumption in the industrial sector. Electric motors are used in most industrial processes and account for more than two-thirds of electricity consumption in this sector [1].

Wireless networks present a number of advantages compared to wired networks as, for example, the ease and speed of deployment and maintenance, in addition to their low cost. Wireless Sensor Networks (WSNs) extend the advantages with their self-organization and local processing capabilities. Therefore, WSNs appear as a flexible and inexpensive solution to build industrial monitoring and control systems [2, 3].

Mainly due to high cost of monitoring systems, in general only motors over 500 HP are monitored. However, motors that have less capacity account for 99.7% of the motors in operation, accounting for about 71% of power consumption [4].

Among the parameters that can be monitored, torque is crucial to prevent motor failures, avoiding losses in the production process [5]. The shaft torque can be estimated from the motor electric signals. Although this technique is less accurate when compared to the direct measurement methods, it is less invasive and allows shaft torque monitoring

using low cost voltage and current sensors that can be easily integrated into a WSN [6]. The shaft torque can also be used for estimating motor efficiency, which is also the most important factor when targeting power consumption reduction [7, 8].

Other methods for fault analysis in motors may also be employed, like methods based on electrical signature analysis, which verify variations in voltage and current signals, in order to relate the signature characteristics with electrical and mechanical conditions [9–11] or methods based on vibration analysis using accelerometers, which are based on parameters such as displacement, velocity, and acceleration for detecting faults in motors [12, 13].

However, employing WSNs in automation systems in industrial environments presents a number of challenging aspects. Wireless networks have unreliable communication links [14], which can be worsened due to noise and interference in the communication spectrum range. The unreliability of the transmission medium in wireless networks makes it difficult to define quality of service guarantees.

Furthermore, sensor nodes must have low cost, which results in a number of constraints such as low bandwidth and low processing power. For instance, the IEEE 802.15.4 standard presents a nominal throughput of 250 kbps. However, even considering this limitation, such standard can possibly comply to the requirements of many industrial monitoring systems.

Industrial monitoring systems need to measure signals that change rapidly, in a dynamic manner [15]. Applications such as efficiency monitoring and fault detection in induction motors fit this type of application. Due to the limitations of WSNs, mainly regarding the low bandwidth and the lack of reliability in the transmissions, the implementation of such systems becomes even more challenging.

In this context, this paper aims at analyzing the impact of using smart nodes for motor monitoring applications in industrial environments employing WSNs, through theoretical and experimental studies. By performing local processing, it is possible to mitigate the intrinsic limitations of WSNs. However, the processing latency should be considered, especially because the sensor nodes usually have low processing power and reduced memory.

Mathematical models for estimating the information delivery rate in several scenarios are properly described, and the results obtained from the models were analyzed together with experimental results. To perform this analysis, we take into account applications for torque and efficiency monitoring in induction motors. This type of motor corresponds to about 90% among all motors employed in the industry [1]. From these studies, it can be observed that the use of embedded processing in WSNs' nodes allows the monitoring even in high interference scenarios, whereas it would be impractical without local processing.

The proposed mathematical model can be used to assess the impact of using local processing with respect to the information delivery rate of different applications. Some factors affect this metric as, for example, processing delay, the amount of data gathered from sensors, and the packet error rate. Thus, it may be advantageous or not to use local processing, depending on these parameters. Specifically, for the motor monitoring applications analyzed in this paper, it is shown that local processing effectively improves the WSN performance, with a direct impact on the application itself.

2. Industrial Wireless Sensor Networks

In an industrial WSN, sensor nodes are deployed in machinery for monitoring critical parameters such as vibration, temperature, pressure, and efficiency. Measurements are transmitted wirelessly to a sink node, which later provides the gathered information for analysis in a central station. Based on this information, it is possible to repair or replace devices before major damages take place [16].

Although WSNs have several advantages, the deployment of this technology presents some challenges. Wireless communication is inherently unreliable and is subject to a larger number of transmission errors when compared to wired networks, mainly due to channel failures and interference. Nodes can suffer interference from the coexistence with other nodes in the network, from the coexistence with other networks, and from other technologies operating in the same frequency range.

In industrial environments, there may be other sources of interference such as thermal noise, interference from motors and devices that cause electromagnetic interference in the band used for communication [16]. Besides, in industrial environments there is usually a large amount of metallic objects, which can impair the communication. In general, industrial wireless systems are prone to high error rates and often with high variance [17].

2.1. IEEE 802.15.4 Standard. The IEEE 802.15.4 standard was designed for WSN applications. This standard provides wireless communication with low power consumption and low cost for monitoring and control applications that do not require high bandwidth. Compared to other standards, such as IEEE 802.11 (WiFi) and IEEE 802.15.1 (Bluetooth), the IEEE 802.15.4 standard has advantages related to energy consumption, scalability, reduced time for node inclusion, and low cost [18].

The IEEE 802.15.4 standard defines the physical layer and the MAC layer. It has three frequency ranges (868 MHz, 915 MHz, and 2.4 GHz). As these bands are unlicensed, radios share the communication medium with devices that implement other technologies. For example, the IEEE 802.11 and the IEEE 802.15.4 standards both operate in the 2.4 GHz frequency range. However, as the spectrum is divided into channels, it is possible to multiple networks operating simultaneously, without interfering much with each other.

The network topology can be organized in three ways: star, mesh, and tree. However, the standard does not define the network layer. There are two types of nodes: full function device (FFD) and reduced function device (RFD). The FFD nodes can act both as network coordinator or end node. The coordinator is responsible, among other functions, for the initialization, address allocation, network maintenance, and the recognition of all other nodes. RFD nodes work

only as end nodes, which are responsible for the functions of sensing or action. FFD nodes can also perform the function of intermediate routers between nodes, without the intervention of the coordinator [6].

The IEEE 802.15.4 radios reach a nominal throughput of 250 kbps when operating in 2.4 GHz band. Nevertheless, there is an underutilization of the channel due to the CSMA/CA mechanism, for each packet awaits at least one backoff period before transmission. Experimental studies by Lee [19] showed that the real maximum throughput is approximately 153 kbps. For many WSN applications, this throughput is enough, but for applications, where sensors acquire a large amount of data in short periods of time, the available throughput may become insufficient. This problem can be mitigated with the addition of local processing in sensor nodes, which reduces the amount of data transmitted over the network.

Some upper layer protocols have been proposed for IEEE 802.15.4 based WSNs. Following, we present a short description for some of such protocols, as well as implementation issues that impact on the communication reliability.

2.1.1. ZigBee. The most employed protocol in WSNs' applications is the ZigBee protocol [12, 18, 20–24]. This protocol has a number of desirable characteristics for WSN applications as, for example, low power consumption and low cost.

ZigBee's network protocol supports the three available topologies (i.e., star, tree, and mesh), allowing the implementation of an ad hoc WSN. When using a mesh topology, the routing process becomes more complex, but the robustness and fault tolerance of the network increase, due to the ability to find and maintain routes.

ZigBee does not implement mechanisms to mitigate the coexistence problem. For instance, it does not switch channels during periods of high contention and interference; instead, it adopts only a low duty-cycle and medium access control algorithms to minimize data losses from packet collisions [25].

2.1.2. MiWi. The MiWi protocol [26], developed by Microchip, is an alternative for small networks with at most 1000 nodes. Theoretically, a ZigBee network can contain up to 65536 nodes although, in practice, it is not recommended having more than 3000 nodes in a single network [23]. Another factor that limits the size of MiWi networks is the maximum number of hops allowed (i.e., four hops).

An interesting feature that sets MiWi apart from ZigBee is MiWi's ability to perform dynamic channel switching. This mechanism, called *Frequency Agility*, is optional and allows moving the network to operate into a different channel once the current operating conditions are not favorable. To set the new channel, a node, called *initiator*, performs an energy scan in all channels, for finding the least busy one. After that, the *initiator* broadcasts a message to all nodes conveying information regarding the new channel. If a node does not receive the broadcast message from the *initiator* (probably due to a transmission failure), it performs a resynchronization after many recurring failed transmissions. Resynchronization consists of scanning all channels to find out the channel currently in use by the network.

Although this mechanism tends to improve communication quality in general, it incurs an overload on *initiators*. The network may spend much time without providing new data, if the *initiators* perform scans very often. Another important factor is the scanning period. In case it is too long, it is possible to obtain greater accuracy in estimating the best channel; however, the network will be idle for too long. On the other hand, if the scanning period is too short, the network spends little time idle, but then it might present lower accuracy in estimating the best available channel.

Although the *Frequency Agility* mechanism is provided by the MiWi suite, there is a strong dependence on the application layer, since the application determines when a scan and a possible channel switch must occur.

2.1.3. WirelessHART. The WirelessHART standard is considered the first open communication standard designed for wireless industrial monitoring and control applications [27]. The other standards, such as ZigBee and Bluetooth, do not completely meet the requirements of industrial applications.

The WirelessHART is based on the physical layer of IEEE 802.15.4 but implements its own link layer. It is based on the 2.4 GHz ISM band but adopting only 15 channels, because channel 26 is not allowed in some countries [28]. Instead of using CSMA/CA as defined by the IEEE 802.15.4 standard, it implements an MAC layer with Time Division Multiple Access (TDMA). By using TDMA, it reduces collisions and power consumption [29].

To improve the coexistence with other networks and other technologies based on the 2.4 GHz band, the standard implements a frequency hopping mechanism. Another mechanism, called Blacklisting, is also defined. Using this mechanism, the channels with high level of interference are avoided. However, the blacklist is not done automatically but by a network administrator.

In a WirelessHART network all nodes on the network must be able to perform routing. It is used a mesh topology with redundant routes. This feature allows increasing the reliability and fault tolerance, since redundant routes can replace obstructed paths. The routes are generated by a central entity (network manager). The network manager is also responsible for scheduling time among the nodes of the network, ensuring the correct operation of the TDMA mechanism.

WirelessHART networks are centralized, because the entire network operation is managed by a single entity. In MiWi or ZigBee networks, end nodes discover their route to the destination. Moreover, each node can decide when to initiate a transmission independently, using the CSMA/CA mechanism. In WirelessHART, the network manager defines the moment when each node should transmit or receive packets.

Petersen and Carlsen [28] performed studies on the performance of WirelessHART radios. The performance of these radios was also verified when subject to interference from three IEEE 802.11 g access points (operating in channels

1, 6, and 11). The results showed that, during interference periods, nodes experienced an average packet error rate of 27.2%. However, in the experiment all channels were enabled. With a better management of blacklists, the performance could be improved. However, radios do not have the ability to enable and disable channels automatically.

A large latency of around two seconds for the network operating without interference and around 2.7 seconds when operating in coexistence with the access points was observed. Thus, we can see that the overhead due to the mechanisms implemented to increase reliability implies a large latency.

WirelessHART is a recent standard, released in 2007. Until 2009 there was no complying component available on the market [29]. However, more experimental studies should be conducted to verify the performance of WSN that comply with this standard.

2.1.4. ISA100. The Instrumentation, Systems, and Automation Society (ISA) idealized the ISA100 standard [30], which is also designated for industry. As the WirelessHART, the ISA standard is based on the IEEE 802.15.4 physical layer but defines its own MAC layer. The MAC layer characteristics are very similar to the characteristics presented on WirelessHART. It also applies TDMA and frequency hopping to improve reliability. The network layer is a bit different, since it uses header formats based on the IP protocol [27].

2.1.5. Comparison among the Standards. Table 1 presents a brief comparison among the standards under consideration with respect to some aspects.

ZigBee is the only protocol that presents no special mechanism for coexistence. The MiWi protocol provides a mechanism for switching channels, but there is still much dependence on the application layer.

On the other hand, the WirelessHART and ISA100 standards offer more complex mechanisms to improve the coexistence for industrial WSN. The main drawbacks are the heavy network centralization and the high communication latency, which results in a low information delivery rate [31]. Furthermore, from [28] we can see that, if there is no proper blacklist management, network performance can suffer a significant drop in the presence of interference.

WirelessHART and ISA100 also implement redundant routes, which can increase the reliability, since multiple paths may be defined for data transfer. However, as this mechanism is implemented at the network layer, it can also be implemented in radios that comply with the physical and MAC layers of IEEE 802.15.4.

Although WirelessHART and ISA100 are intended for industrial WSN applications, these are pretty new standards, and they do not have high availability of complying transceivers on the market. On the other hand, there is a wide availability of transceivers that implement the physical and MAC layers of IEEE 802.15.4 and are compatible with ZigBee and MiWi.

The results of this paper consider radios fully compatible with IEEE 802.15.4. However, it is important to notice that they are totally applicable to other protocols. The use of local processing still remains very important, due to the low throughput of the IEEE 802.15.4 physical layer. In addition to that, WirelessHART and ISA100 present reliability concerns [28].

3. Motor Monitoring Systems

This section describes some works [6–8, 15, 32–35] focusing on the application of WSN in industrial environments. There is a relatively small amount of work towards the development of monitoring and control systems in industry based on WSNs. This is due to the complex requirements of the system and severe work environment [15]. Some recent works address the performance evaluation of radios in an industrial environment [2, 25, 36–38], while some other works address the challenges of using WSN technology in industry [16, 17, 39–41].

Salvadori et al. [32] proposed a digital system for evaluation of power usage, diagnosis, control, and supervision of electrical systems employing WSNs. The system is based on two hardware topologies responsible for signal acquisition, processing, and transmission: intelligent sensor modules (ISMs) and remote data acquisition units (RDAUs). However, only wired communication RDAUs are used to perform acquisition of voltage and current of motors. ISMs were used only for temperature measurements. The work focuses mainly on the energy consumption of sensor nodes and does not provide detailed studies on transmission errors and communication channel quality.

Hsu and Scoggins [42] presented a method to estimate motor efficiency from the air-gap torque, which is obtained from the motor electrical signals (current and voltage). It is the noninvasive method for determining torque and efficiency that has less uncertainty [43]. Recent works have also used this technique to estimate the efficiency and torque of induction motors [6–8, 35]. These studies have also employed WSN for data transmission.

Lu et al. [7, 8] identify in their work the synergies between WSNs and analysis of motors based on electrical signals, also following a noninvasive pattern. They propose a scheme to apply WSNs for online and remote monitoring and fault diagnosis of industrial motors.

The main limitation of the work presented in [7, 8] is derived from the low throughput provided by the WSN based on the IEEE 802.15.4, since the proposed system does not employ local processing. Thus, it is necessary to transmit a large amount of data to estimate the desired parameters. This limits, among other things, the acquisition rate from the sensors, which consequently limits the accuracy of the estimation. In a WSN with a large number of nodes, the situation becomes even worse, since all nodes share the same physical medium. Moreover, the wireless networks are inherently unreliable, which can result in transmission errors, affecting the estimation process.

Hou and Bergmann [33] developed a motor monitoring system using WSN with local processing. A prototype was implemented and validated in a single-phase induction motor in laboratory. Motor current signature analysis (MCSA) is

Table 1: Comparison among the standards.

Standard	ZigBee	MiWi	WirelessHART/ISA100
Supported frequency bands	868 MHz, 915 MHz, and 2.4 GHz	868 MHz, 915 MHz, and 2.4 GHz	2.4 GHz
Physical layer	IEEE 802.15.4	IEEE 802.15.4	IEEE 802.15.4
MAC layer	IEEE 802.15.4	IEEE 802.15.4	Custom
Medium access mechanism	CSMA/CA	CSMA/CA	TDMA
Coexistence mechanisms	—	*Frequency Agility*	Frequency hopping and Blacklisting
Definition of routes	Distributed	Distributed	Centralized
Redundant routes	No	No	Yes

employed in this application, where motor stator current signal waveforms are given under different working conditions. Using local processing, a reduction of around 90% was obtained in the amount of data transmitted for analyzing the motor.

Hou and Bergmann [15] also developed a system for fault detection in motors using accelerometers and WSN. In this system, a reduction of 99% was obtained in the amount of data transmitted for performing the failure analysis task when using local processing. However, both [15, 33] do not performed a detailed analysis of the WSN performance. For example, the information delivery rate was not verified, and the experiments were not performed in realistic environments. When employing local processing in WSN systems, it is also important to consider processing latency, as well as the quality of the communication medium, to effectively verify the gain obtained from local processing.

Hu [34, 35] presents a DSP-based system for motor monitoring using the air-gap torque method and WSN for data transmission. The estimation of various parameters such as power factor, efficiency, speed, and torque was proposed. However, the tests were conducted in laboratory, which does not characterize a realistic experiment. As in [15, 33] there was no detailed study on the impact of using local processing on the WSN performance.

In a previous work [6], we developed an embedded system integrated into a WSN for online dynamic torque and efficiency monitoring in induction motors. The air-gap torque method was employed for estimating the shaft torque and motor efficiency. The computations for estimating the targeted metrics are performed locally and then transmitted to a monitoring base unit through an IEEE 802.15.4 WSN.

Experimental tests were performed to analyze the torque values obtained by the system and then compared with torque values based on the workbench dynamic model. The paper also showed an experimental study aiming at identifying the correlation between spectral occupancy and packet error rate (PER) for the proposed WSN. The experiments were conducted inside a shed, with typical characteristics of industrial environments.

The study demonstrated that the addition of new interference sources can significantly affect the spectral occupancy by also having a direct impact on the communication performance. Even for harsh condition scenarios, the system was able to provide useful monitoring information, since all processing is done locally (i.e., only the computed metric is transmitted over the network). Without local processing, it might be impossible to use the WSN technology for this particular application, considering an unreliable transmission medium.

In [6] a theoretical study was conducted to determine the number of packets transmitted over the network, comparing the approaches with and without local processing, as well as with and without packet retransmission. In this paper, we extend the mathematical model described in [6], to verify the information delivery rate in different scenarios, with varying PER and taking into account processing delay. In addition to that, we analyze the impact of retransmissions and aggregation. The results obtained through the analytical model were analyzed together with experimental results to show the impact of using local processing for motor monitoring applications based on WSN in industrial environments.

4. System Description

In this section, we describe the system designed for performing our studies. The system (Figure 1) consists in a WSN running an application for torque and efficiency measurements in induction motors. More details about this system can be found in [6].

End nodes are composed of embedded systems located close to the electric motors. Sensors provide motor voltage and current measurements, and the embedded system performs the required processing for computing torque and efficiency. After local processing, the metrics are transmitted to the base station through the WSN.

Internally, end nodes (Figure 2) are composed of current and voltage sensors and an acquisition and data processing unit (ADPU), which is responsible for data acquisition and A/D conversion, besides data processing. Finally, an IEEE 802.15.4 transceiver is used for communication in the WSN.

Torque and efficiency are computed from the motor electrical signals, using the air-gap torque method, since this is the noninvasive method that presents less uncertainty [43]. To perform the estimation using this method in a three-phase induction motor, it is necessary to acquire two voltage signals and two current signals. More details about the air-gap torque method can be found in [4, 6–8].

Figure 3 shows a flowchart of the embedded system's operation. When the system starts, all embedded system's

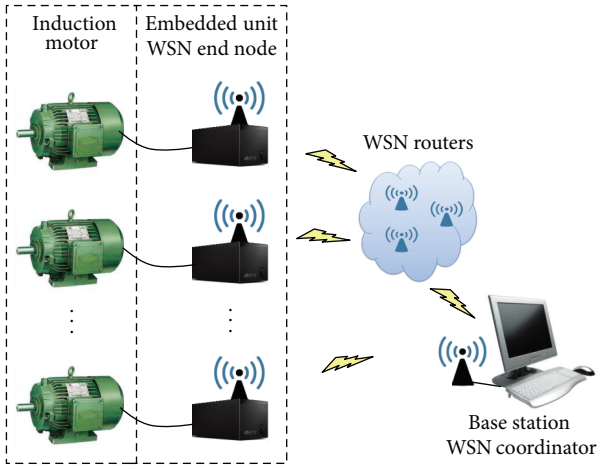


Figure 1: Proposed WSN [6].

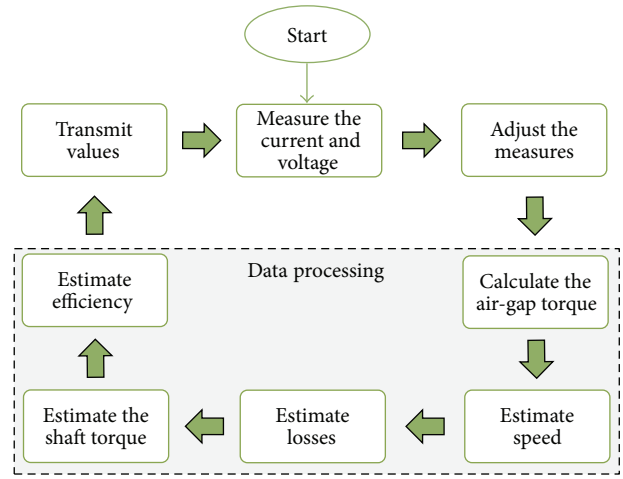


Figure 3: Flowchart of the embedded system.

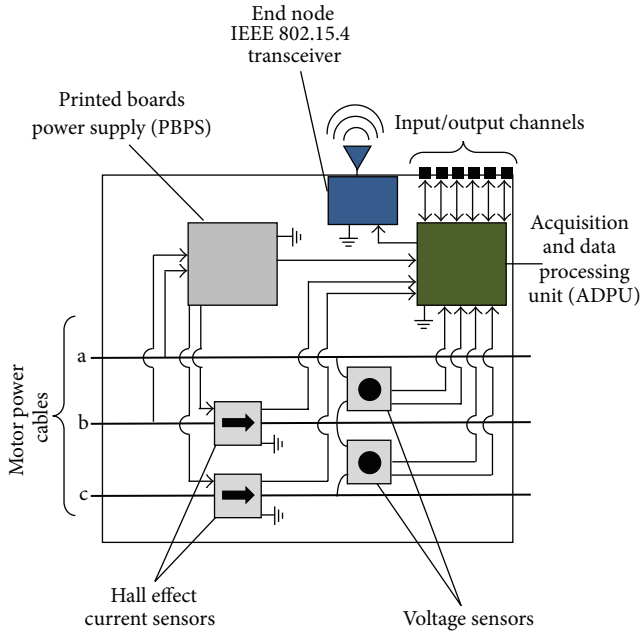


Figure 2: : Embedded system [6].

parameters are configured, such as the A/D converter and the network parameters. Then, the system performs a cycle that includes the acquisition of current and voltage values, sampling shaping (e.g., offset removal), data processing to estimate torque and efficiency, and transmission of the target values through the WSN.

5. Theoretical Analysis

5.1. Information Delivery Rate. In this section we present a model to analyze the latency for information delivery (torque and efficiency values) in the WSN. The propagation time was considered null in the model. Thus, the latency for information delivery by the WSN is computed only in terms of transmission time and processing time. We considered a WSN with star topology, which is the most prevalent topology today [29].

The latency (L) for receiving one torque value and one efficiency value from the WSN, in the scenarios without retransmission, is given by the following equation:

$$L = Q_t \left(Q_p \left(\frac{P_{\text{size}}}{W} + e(p) \right) + L_p \right), \quad (1)$$

where Q_t is the number of trials until information is delivered; Q_p is the number of packets transmitted in each trial; P_{size} is the number of bits of each packet; W is the network bit rate; $e(p)$ is the error due to the assumptions of the model; L_p is the processing latency.

Note that the value of W represents the number of bits transmitted per second, regardless of transmission errors.

The latency (L_r) for receiving one torque value and one efficiency value from the WSN, in the scenarios with retransmission, is given by the following equation:

$$L_r = Q_t \left(Q_p \left(\frac{P_{\text{size}}}{W} + \frac{H_{\text{ack}}}{W} + e(p) \right) \right) + L_p, \quad (2)$$

where H_{ack} is the number of bits of an acknowledge packet (ACK).

5.1.1. Impact of Packet Loss. The transmission of a packet consists on a Bernoulli event with successful probability p, with a number of trials until the first success as defined by a geometric distribution. In a geometric distribution, the average number of events until the first success is $1/p$.

Thus, the probability of successfully transmitting the data necessary for estimating the torque and efficiency is p^{Q_p}, and the number of attempts before the first success is $Q_t = 1/p^{Q_p}$, considering a scenario without retransmission of lost packets. As in each attempt Q_p packets are transmitted, then we have an average of Q_p/p^{Q_p} transmissions per efficiency and torque measures.

Considering that the probability of successfully transmitting an acknowledgment packet is also p, then the probability

of successfully transmitting a packet in the scenario with retransmission is p^2, and the number of trials until the first success is $1/p^2$. Thus, the average number of transmissions until the first hit is $Q_t = Q_p/p^2$.

5.1.2. Scenario without Local Processing. Each node in the WSN obtains data from two current sensors and two voltage sensors. Current and voltage signals have a frequency of 60 Hz, and computing one value of air-gap torque requires a complete cycle of voltage and current. Therefore, the total number of bits for estimating the target values (Q_b) can be defined according to the following equation:

$$Q_b = \frac{4bT_a}{60}, \tag{3}$$

where T_a is the acquisition rate of the analog to digital converter (ADC) and b is the amount of bits per sample (which depends on the resolution of the ADC).

Then, without using local processing, the number of packets, (Q_p), required for transmitting the data acquired by the sensors is given by the following equation:

$$Q_p = \lceil \left(\frac{Q_b}{C_m} \right), \tag{4}$$

where C_m is the maximum payload of a packet (defined by the standard).

The average size of the packet payload, (C_{avg}), is given by the following equation:

$$C_{\text{avg}} = \frac{Q_b}{Q_p}. \tag{5}$$

Thus, in the scenario without local processing $P_{\text{size}} = C_{\text{avg}} + H$, where H is the number of bits of the packet header.

In the scenario without local processing, only the time to acquire the values, equal to 1/60 seconds, was considered as the total processing latency; that is, $L_p = 1/60$.

As demonstrated in Section 5.1.1, $Q_t = 1/p^{Q_p}$ for the scenario without retransmission, and $Q_t = 1/p^2$ for the scenario with retransmission.

Therefore, the latency for receiving one torque value and one efficiency value from the WSN, in the scenario without local processing and without retransmission ($L_{n\text{lp}}$), is given by the following equation:

$$L_{n\text{lp}} = \frac{1}{p^{Q_p}} \left(Q_p \left(\frac{C_{\text{avg}} + H}{W} + e(p) \right) + \frac{1}{60} \right). \tag{6}$$

The latency for receiving one torque value and one efficiency value, in the scenario without local processing and with retransmission ($L_{r_{n\text{lp}}}$), is given by the following equation:

$$L_{r_{n\text{lp}}} = \frac{1}{p^2} \left(Q_p \left(\frac{C_{\text{avg}} + H}{W} + \frac{H_{\text{ack}}}{W} + e(p) \right) \right) + \frac{1}{60}. \tag{7}$$

5.1.3. Scenario with Local Processing. With local processing, the nodes acquire Q_b bits to estimate torque and efficiency, but it is necessary to transmit only a constant number of bits (V_s), regardless of the values of b and T_A. With local processing, only one packet is needed to transmit the information; that is, $Q_p = 1$. Additionally, it is possible to aggregate various torques and efficiency values into a single packet.

Let V_a be the amount of values aggregated into the same packet; we have that $P_{\text{size}} = V_a V_s + H$, where $V_a V_s \leq C_m$.

Let P_t be the processing time to estimate the values of torque and efficiency; we have that $L_p = V_a * (P_t + (1/60))$.

Since $Q_p = 1$, we have that $Q_t = 1/p$ for the scenario without retransmission, and $Q_t = 1/p^2$ for the scenario with retransmission.

Therefore, the average latency for receiving one packet from the WSN in the scenario with local processing and without retransmission (L_{lp}) is given by the following equation:

$$L_{\text{lp}} = \frac{1}{p} \left(\frac{V_a V_s + H}{W} + e(p) + V_a \left(P_t + \frac{1}{60} \right) \right). \tag{8}$$

In the scenario with retransmission and with local processing, the average latency ($L_{r_{\text{lp}}}$) for receiving one packet from the WSN is defined according to the following equation:

$$L_{r_{\text{lp}}} = \frac{1}{p^2} \left(\frac{V_a V_s + H}{W} + \frac{H_{\text{ack}}}{W} + e(p) \right) + V_a \left(P_t + \frac{1}{60} \right), \tag{9}$$

where $e(p)$ is the error on the estimation of the transmission time and processing time, due to the assumptions. The mathematical model does not take into account propagation delay and the processing overhead regarding the MAC protocol, among other factors.

5.1.4. Error Estimation. The mathematical model described here still presents some limitations, due to the assumptions made. In the scenarios with retransmission, the processing time from the acknowledge mechanism and the timeout for lost packets was also not considered.

It is also important to consider some relations, which are not directly supported by the models. Although the value of W is not directly affected by the value of p, since W is the number of bits transmitted per second, regardless of transmission errors, there is some correlation between these parameters. W is affected by the CSMA/CA mechanism, and when the interference level in the environment is high or the network has many nodes, the packet error rate is likely high. At the same time, there is an increase in backoff periods, having a direct impact on W. Therefore, some scenarios are very unlikely to happen as, for example, the scenario with $p = 0.1$ and $W = 150$ kbps (maximum real throughput).

The relation between packet size and packet error rate was not taken into account in the analytical model. Thus, in practice the WSN performance will be lower than the one computed through the models. However, the model is useful for comparing the performance of several possible

approaches. Particularly, it can be observed that the use of local processing can significantly increase the WSN performance. It should also be noted that when local processing is not used, the real performance suffers even more with the assumptions made, because there are more and larger packets being transmitted over the network.

To verify the precision of the models, some tests were performed using the embedded system developed (described in Section 4) and a coordinator node, to measure the information delivery rate and compare the results with the results obtained using the models. It is important to note that only some scenarios were analyzed. It was not possible to validate the scenarios with retransmission, since the radios used in our testbed do not provide information about transmission errors. Thus, it was not possible to identify the packet error rate. When retransmission is not used, to obtain the packet error rate, it is only necessary to verify the number of packets transmitted and the number of packets received at the upper layer.

During these experiments, the radios were configured to operate on channel 11. First, we obtained results for the scenario without packet loss, given that there was no interference source in the environment. To obtain values for the scenarios with packet loss, we activate a WiFi network in the environment, operating on channel 1, which coexists with channel 11 of IEEE 802.15.4 in the 2.4 GHz band. Although no significant traffic has been injected in the WiFi network, it was possible to observe a drop in the communication performance of the IEEE 802.15.4 radios. The embedded system was configured to transmit 1000 packets, and the coordinator was instrumented to calculate the information delivery rate. The time for receiving the packets was also measured by the coordinator.

With the results obtained from the experiments, it was possible to compute the values of $e(p)$ for the experiments under consideration. As discussed earlier, there is a correlation between the values of p and W. $e(p)$ corresponds to the additional time of transmission and processing overhead related to the errors in the models. Therefore, the value of $e(p)$ must compensate these assumptions. Thus, a model was defined to estimate these values for each scenario from the value of p, since the smaller the value of p, the greater the error.

To relate these two parameters a nonlinear regression (exponential) was performed. Through these regressions, we obtained (10), (11), (12). Since the error is different in each scenario, a model for each one was obtained

$$e_{n\text{lp}}(p) = 0.0073(0.202)^{p}, \tag{10}$$

$$e_{\text{lp}}(p)_1 = 0.0061(0.363)^{p}, \tag{11}$$

$$e_{\text{lp}}(p)_{18} = 0.029(0.115)^{p}, \tag{12}$$

where $e_{n\text{lp}}(p)$ is the error for the scenario without local processing. $e_{\text{lp}}(p)_1$ is the error for the scenario with local processing, and $V_a = 1$. $e_{\text{lp}}(p)_{18}$ is the error for the scenario with local processing, and $V_a = 18$.

In the next section, we will analyze the model accuracy, compared with values obtained by the experiments.

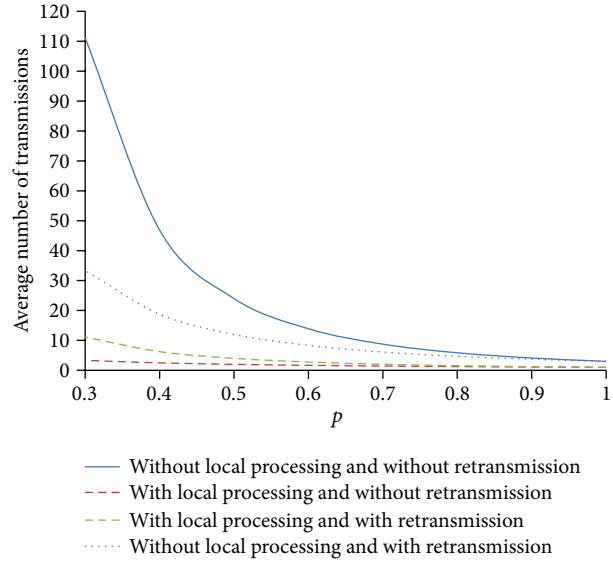


Figure 4: Average number of transmissions.

6. Results

Using the generic mathematical model described above, we will conduct an analysis considering the IEEE 802.15.4 characteristics. In this standard, $C_m = 118$ bytes. 4 bytes are necessary to store one torque value and one efficiency value (i.e., $V_s = 32$ bits). To obtain good accuracy from a simple numerical integration method, such as trapezoidal (the one used to implement the algorithm), it requires a sample rate greater than 2 kHz [4]. Based on these observations, T_a was set to 3 KHz, and $b = 10$ bits.

6.1. Number of Transmitted Packets. Figure 4 shows the number of packets that must be transmitted, on average, for the four scenarios. In the chart, the values for $p = 0.2$ and $p = 0.1$ were omitted, in order to improve the visualization. For $p = 0.1$, without local processing and without retransmission, we have 3000 transmissions, while in the approach with local processing and without retransmission only 10 transmissions are required on average. In the chart, we counted only the data transmissions; that is, the transmission of the ACK packets was not taken into account.

It is important to notice that as torque and efficiency values occupy only 4 bytes, we can aggregate multiple values in a single packet. Besides that, due to an increase in the number of transmitted packets, considering an approach without local processing, the packet error rate tends to increase; that is, the value of p tends to be lower.

6.2. Latency. At first, we considered $e(p) = 0$ in all scenarios. Figures 5 and 6 show the delivery latency for one torque value and one efficiency value, for the scenarios with local processing and without local processing, respectively. In Figure 6 $V_a = 1$ and $L_p = 14$ ms. The other parameters were defined according to the IEEE 802.15.4 standard. We conducted a performance test with the embedded system

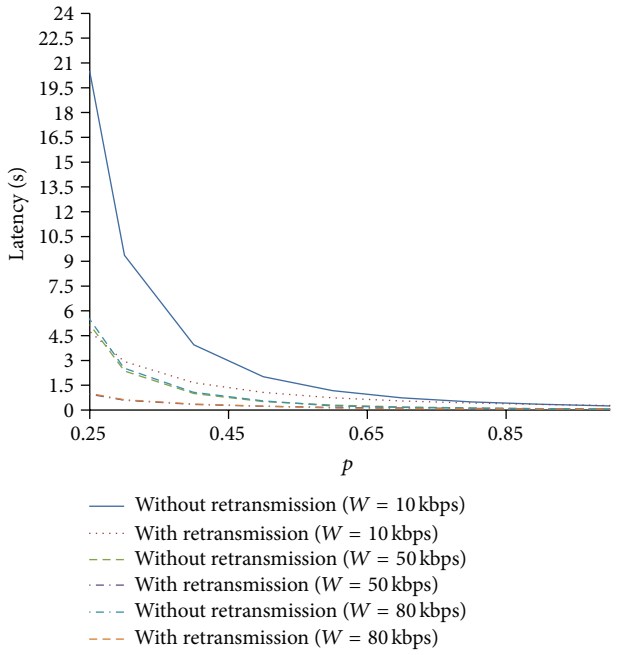


Figure 5: Latency to receive information in the WSN (without local processing).

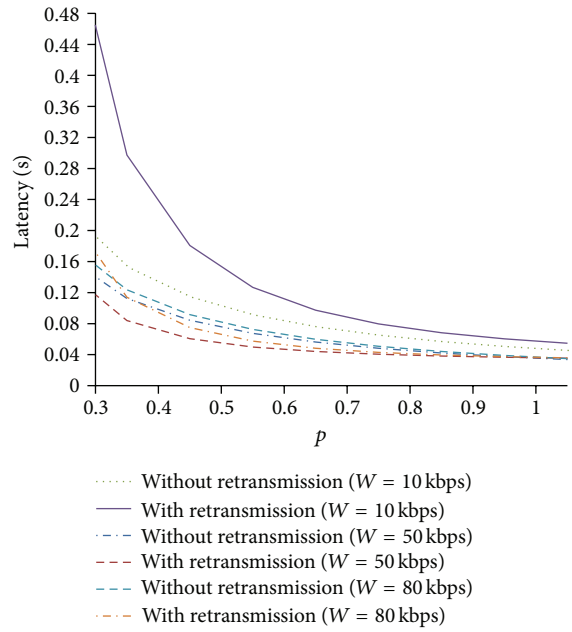


Figure 6: Latency to receive information in the WSN (with local processing).

implemented in this work (considering $T_a = 3000$ Hz) to observe the processing time for the metrics under consideration. The results pointed out a processing time of around 14 ms.

Latency values are shown for various values of p and for different bit rates (10 kbps, 50 kbps, and 80 kbps). The maximum effective bit rate for the IEEE 802.15.4 standard is around 150 kbps, but this rate may be much lower with an increasing number of nodes in the network, leading to longer backoff periods before nodes perform packet transmissions. For example, in [19] it was demonstrated that the data rate in a WSN with four nodes reaches a maximum of around 80 kbps.

From the charts, we can see that when local processing is used, latency is short, even in scenarios with high PER and low bit rate. The charts do not include values for $p < 0.25$, to enhance visualization. Without local processing and without retransmission, latency can reach up to 253 seconds for $p = 0.1$ and $W = 10$ kbps. Also for $p = 0.1$, latency can be as low as 2.93 seconds, for $W = 80$ kbps and with retransmission.

With local processing, in the worst scenario (i.e., $p = 0.1$ and $W = 10$ kbps), latency is around 2.43 seconds when using retransmission. We can notice, from Figures 5 and 6, that retransmissions improve the overall performance in scenarios without local processing for all considered bit rates. This is due to the large amount of transmissions and minimum processing time. When using local processing, retransmissions may be advantageous only in some cases, when both throughput and p are high.

6.3. Information Delivery Rate. If many values of torque and efficiency are aggregated into a single packet, latency increases due to an increase in the overall processing time. However, the information delivery rate may be larger, because each packet carries more information. Therefore, to verify the WSN performance when using aggregation, we will analyze the information delivery rate, that is, the amount of torque and efficiency information received per second for a specific node (the inverse of latency).

The chart in Figure 7 shows a comparison between the values obtained from the experiments (reference) and the values obtained from the model (8), for the scenario with local processing, and $V_a = 1$. Three curves are plotted from the model, besides the reference curve.

Since it is not possible to observe with accuracy the value of W, the two first curves were obtained from the model for $W = 150$ kbps and $W = 80$ kbps, respectively, and $e_{lp}(p)_1 = 0$. Also, the absolute errors between the curves obtained from the model and the reference curve, for the three cases are shown.

The third curve obtained from the model considers the value of $e_{lp}(p)_1$, obtained from (11). To obtain the model of $e_{lp}(p)_1$, the value of W was set at 80 kbps. Although the value of W can vary depending on the value of p, the error resulting from the constant value of W is also compensated by $e_{lp}(p)_1$.

For $W = 80$ kbps, the errors vary from 6.2% to 8.8%, and for $W = 150$ kbps the errors vary from 8.7% to 11.2%. We can notice that the smaller the value of p, the greater the error between the reference and the model. This occurs due to the relation between p and W.

When $e_{lp}(p)_1$ is considered, the error stays less than 1% for almost all values of p, reaching a maximum of 1.41%, as observed in Figure 7.

The chart in Figure 8 shows a comparison between the values obtained from the experiments (reference) and the

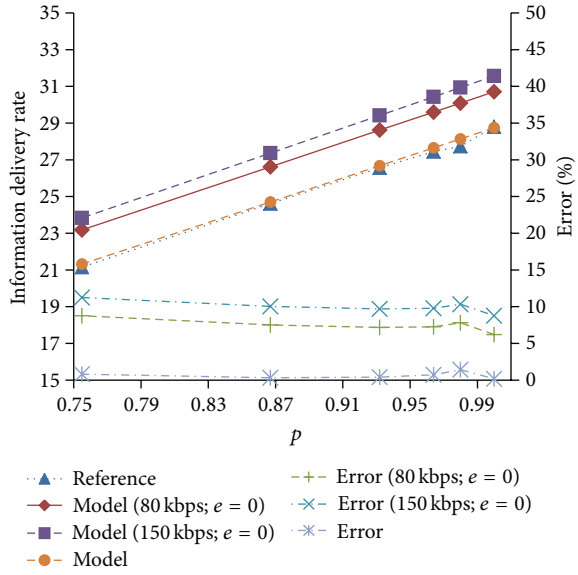


Figure 7: Validation of the model for the scenario with local processing, without retransmission, and $V_a = 1$.

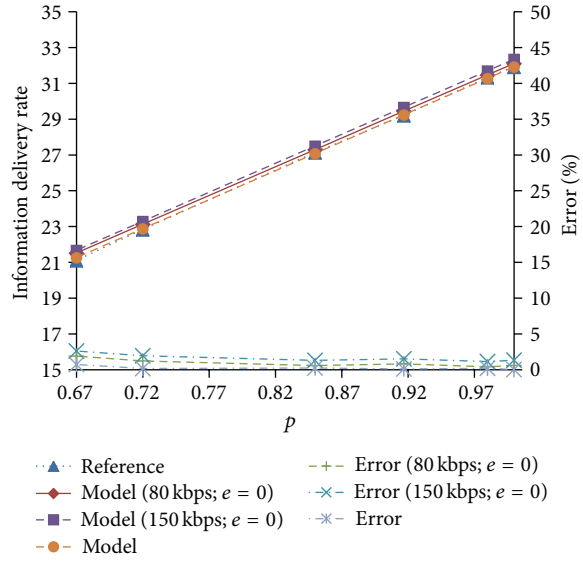


Figure 8: Validation of the model for the scenario with local processing, without retransmission, and $V_a = 18$.

values obtained from the model, for the scenario with local processing, without retransmission, and $V_a = 18$.

When $W = 80$ kbps, the errors vary from 0.6% to 2.36%. When $W = 150$ kbps, the errors vary from 1.32% to 3.07%. Again, we can observe that the error increases when p decreases. The overall error was smaller in this scenario, even considering $e_{\text{lp}}(p)_{18} = 0$, because the system spends more time performing processing to compute the torque and efficiency values, spending less time transmitting, since each packet carries 18 values of torque and 18 values of efficiency. Since the main error source in the analytical model comes from transmission time, neglecting it does not have a larger impact on this particular scenario.

When assuming $e_{\text{lp}}(p)_{18}$, the error reaches a maximum of 0.72%, as observed in Figure 8.

The chart in Figure 9 shows a comparison between the values obtained from the experiments (reference) and the values obtained from the model, for the scenario without local processing and without retransmission.

When $W = 80$ kbps, the error is up to 13.33%, and when $W = 150$ kbps, the error is up to 39.17%. In this scenario, the error was larger for $e_{\text{nlp}}(p) = 0$, because the system spends most of the time transmitting, and the processing latency is minimal, making the transmission time the main error source. However, when $e_{\text{nlp}}(p)$ is considered, the model presents a small error of up to a 3.99%, even with the great number of transmissions in this scenario.

Figure 10 shows the information delivery rate (obtained by the models) with p varying from 0.1 to 1, considering the values of $e_{\text{nlp}}(p)$, $e_{\text{lp}}(p)_1$, and $e_{\text{lp}}(p)_{18}$.

As it was not possible to validate the scenarios with retransmission, for this analysis we considered that the errors in the scenarios with retransmission are equal to the errors when retransmission is not used. However, the error values

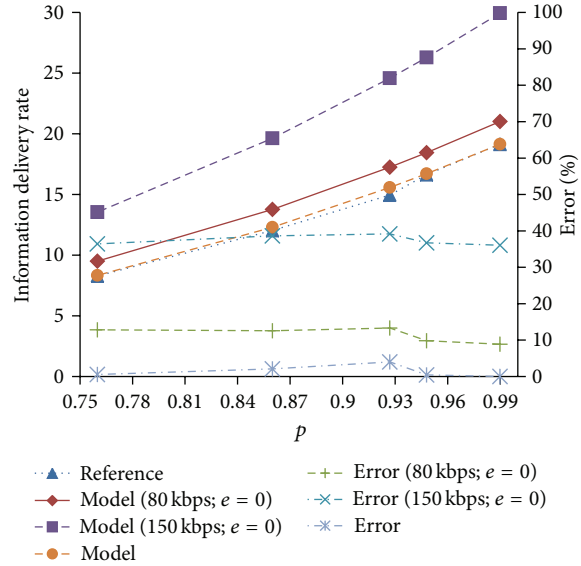


Figure 9: Validation of the model for the scenario without local processing and without retransmission.

for the scenarios with retransmission are larger, due to the assumptions made regarding the ACK mechanism.

We can conclude that when employing local processing and for $V_a = 18$, the information delivery rate is better for all values of p. Without local processing the information delivery rate tends quickly to zero when p approaches zero. Thus, monitoring may be infeasible in environments with a lot of interference and without local processing.

For $V_a = 1$ we believe that the performance is worse when retransmission is used. Although for some values of p the information delivery rate is larger, the model is very simplistic in the modeling of the time associated with the

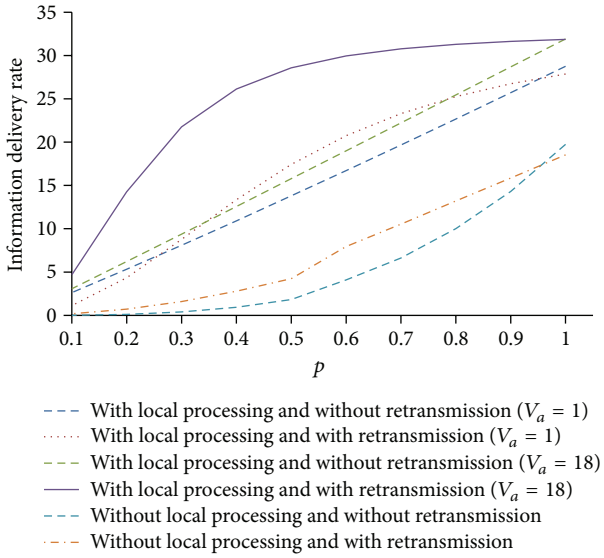


Figure 10: Information delivery rate of the WSN, considering the values of e.

ACK mechanism, and the real performance is worse than the one observed in the chart.

The use of retransmission indicates an improvement in the information delivery rate for $V_a = 18$. Although the real performance is less for the scenario with retransmission than the observed in the chart, we believe that for a large range of p the use of retransmission can be advantageous. This is due to the processing time to calculate 18 values of torque and 18 values of efficiency.

To verify the WSN performance when using retransmission, more experimental studies will be necessary. From these studies, it is also possible to refine the model for the scenarios with retransmission.

7. Discussion

In scenarios with high level of interference, the use of local processing becomes very important. Many recent works address the performance evaluation of radios in an industrial environment [2, 25, 36–38], when subject to interference.

In our previous work [6] a study to verify the impact of interference sources (microwave oven and IEEE 802.11g network) in the communication quality of the WSN described in Section 4 in an industrial environment was performed. We will relate these results with the theoretical results obtained in this paper.

Based on the theoretical study in Section 5, Table 2 shows the information delivery rate for some values of p.

We can observe once more that when p is small, the information delivery rate is very small when local processing is not used. On the other hand, when using local processing, it is possible to perform monitoring even with high packet error rate.

From the studies in [6], we observed that when the WSN (operating on channel 18) was exposed to the interference of an IEEE 802.11g network operating in channel 6, the packet error rate reached about 90%. In this case, the latency to transmit an information (values of torque and efficiency) can reach 68 seconds, when both local processing and retransmission are not used, and at least 5.5 seconds when retransmission is used. Thus, in the best case, nodes deliver 0.18 values of torque and efficiency per second. Moreover, the real information delivery rate is less than that, due to the overhead of the ACK mechanism, which was not considered completely in the model.

For a packet error rate of 90% ($p = 0.1$), when local processing is used, and $V_a = 18$, in the worst case (without retransmission) the latency to transmit an information is about 0.32 seconds; that is, it is possible to obtain about 3 values of torque and efficiency per second from the nodes, even for this high interference scenario.

From these studies we can notice that the deployment of a WSN in industrial environment still presents serious challenges related to the communication reliability. However, despite the high packet error rate in some cases, it is important to notice that, due to the local processing capability, all packets that are received in the destination carry useful information, even considering the processing latency by the embedded system. Without local processing, it is necessary to transmit many packets to provide the desired information, and it is practically impossible to obtain useful data from the WSN without using local processing, for scenarios with high interference.

Besides the application for torque and efficiency monitoring, other applications may suffer even more with the unreliability and low bandwidth of WSNs. For example, in [7] the motor current signature analysis method for fault detection was employed. The ADC was configured to operate with an acquisition rate of 4 kHz and 12 bits of resolution. To perform the analysis, it was necessary to acquire values during 10 cycles of current, which results in about 8000 bits of data (i.e., ($Q_b = 8000$)). Thus, nine IEEE 802.15.4 packets are required to store all the data. For example, for $p = 0.2$, without local processing and without retransmission, we have 17578125 transmissions in average, and with retransmission 225 transmissions in average (data transmission plus acknowledgment) are necessary. When using local processing and without retransmission, on average only 5 transmissions are needed. However, the processing latency of the fault detection algorithm must be analyzed, to verify the real gain obtained with local processing.

8. Conclusions and Future Work

In this paper, we have presented a theoretical study for verifying the performance of motor monitoring systems in industry employing WSN. First, a discussion about the standards and protocols already proposed for WSN and about some implementation aspects which can impact the quality of service in WSN based applications in industrial environments was performed. Finally, mathematical models were developed for verifying the performance of an IEEE 802.15.4 based WSN for applications of torque and efficiency monitoring

Table 2: Information delivery rate.

	LP ($V_a = 1$)	LPR ($V_a = 1$)	LP ($V_a = 18$)	LPR ($V_a = 18$)	NLP	NLPR
$p = 0.1$	2.63	1.13	3.08	4.69	0.015	0.18
$p = 0.2$	5.33	4.34	6.22	14.27	0.12	0.72
$p = 0.3$	8.09	8.77	9.38	21.76	0.4	1.6
$p = 0.4$	10.92	13.36	12.57	26.14	0.94	2.79
$p = 0.5$	13.8	17.43	15.78	28.57	1.84	4.25

LP: with local processing.
LPR: with local processing and with retransmission.
NLPR: without local processing and with retransmission.
NLP: without local processing.

in induction motors, which are widely used in industries. Methods for estimating torque and efficiency in a noninvasive manner have been studied for years, and some works [6–8, 35] have already proposed the integration of these methods with the WSN technology.

From the models, it was possible to verify the performance of the WSN in several scenarios and the benefits from performing local processing, taking into account the lack of reliability of the transmission medium, the characteristics of the IEEE 802.15.4 standard, and processing latency. The analytical results were validated and analyzed together with experimental results obtained in a previous work. It was shown that the use of local processing is essential for the application under consideration, mainly when the WSN is subject to interference sources. In scenarios with high level of interference, it can be almost impossible to perform monitoring without using local processing.

In other applications, such as the fault detection from motor current signature analysis, the use of local processing can be even more essential, since the amount of data that must be processed to perform the analysis is even higher in comparison with the estimation of torque and efficiency. Besides, by performing local processing, the end nodes can be configured to transmit data only if an important event (i.e., a failure) is detected. However, the processing latency to perform the fault detection is also high. Thus, it is important to analyze the performance of the WSN taking into account the node's processing capacity too. The models developed in this paper can be instantiated to evaluate the performance of other applications, such as the fault detection application, verifying the impact of local processing. As one of the future works, we intend to develop fault detection applications using WSN.

Using the results obtained from the model, it is possible to compare several configurations, which can guide de development of WSN applications. This kind of study is especially useful for industrial applications, due to the lack of reliability of wireless networks in this particular environment, and for any application that needs to acquire a large amount of data from sensors.

As future work, we intend to perform detailed performance studies using a WSN with a large number of nodes inside an industrial environment. From these studies we will verify the scalability of the system and characterize the main interference sources in the environment. We also intend to develop spectrum-aware protocols, in which radios can choose the operating channel dynamically, allowing embedded systems to self-adapt to the environment and improving the quality of service of the network. Through more detailed experimental studies, it will be possible to refine the analytical models, increasing their accuracy. Some topics of interest include

(i) experimental evaluation of the WSN with an increasing number of nodes in an industrial environment;

(ii) development and evaluation of protocols for dynamic channel allocation;

(iii) exploration frequency redundancy, employing multiple transceivers;

(iv) experimental evaluation of the WSN using retransmission;

(v) development of techniques for data summarization, reducing even more the amount of packets transmitted in the network.

Acknowledgments

This work was supported by the National Council of Scientific and Technological Development (CNPq, Brazil), by the Coordination of Improvement of Higher Education Personnel (CAPES, Brazil) and by Eletrobras—CEAL.

References

[1] R. Hanitsch, *Energy Efficient Electric Motors*, RIO 02-World Climate & Energy Event, 2002.

[2] R. D. Gomes, M. A. Spohn, A. C. Lima-Filho, E. G. Anjos, and F. A. Belo, "Correlation between spectral occupancy and packet error rate in IEEE 802.15.4-based industrial wireless sensor networks," *IEEE Latin America Transactions*, vol. 10, no. 1, pp. 1312–1318, 2012.

[3] P. Baronti, P. Pillai, V. W. C. Chook, S. Chessa, A. Gotta, and Y. F. Hu, "Wireless sensor networks: a survey on the state of the art and the 802.15.4 and ZigBee Standards," *Computer Communications*, vol. 30, no. 7, pp. 1655–1695, 2007.

[4] B. Lu, T. G. Habetler, and R. G. A. Harley, "Survey of efficiency-estimation methods for in-service induction Motors," *IEEE Transactions on Industry Applications*, vol. 42, no. 4, pp. 924–933, 2006.

[5] A. C. Lima-Filho, F. A. Belo, and R. D. Gomes, "Tests prove out self-powered, wireless, pump torquemeter," *Oil and Gas Journal*, vol. 106, no. 46, pp. 43–48, 2008.

[6] A. C. Lima-Filho, R. D. Gomes, M. O. Adissi, T. A. B. Silva, F. A. Belo, and M. A. Spohn, "Embedded system integrated into a wireless sensor network for online dynamic torque and efficiency monitoring in induction motors," *IEEE/ASME Transactions on Mechatronics*, vol. 17, pp. 404–414, 2012.

[7] B. Lu and V. C. Gungor, "Online and remote motor energy monitoring and fault diagnostics using wireless sensor networks," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 11, pp. 4651–4659, 2009.

[8] M. H. Benbouzid, "A review of induction motors signature analysis as a medium for faults detection," *IEEE Transactions on Industrial Electronics*, vol. 47, pp. 984–993, 2000.

[9] A. Gandhi, T. Corrigan, and L. Parsa, "Recent advances in modeling and online detection of stator interturn faults in electrical motors," *IEEE Transactions on Industrial Electronics*, vol. 58, pp. 1564–1575, 2011.

[10] J. Antonino-Daviu, S. Aviyente, and E. Strangas, "Scale invariant feature extraction algorithm for the automatic diagnosis of rotor asymmetries in induction motors," *IEEE Transactions on Industrial Informatics*, vol. 9, pp. 100–108, 2013.

[11] A. Bouzida, O. Touhami, R. Ibtioen, A. Belouchrani, M. Fadel, and A. Rezzoug, "Fault diagnosis in industrial induction machines through discrete wavelet transform," *IEEE Transactions on Industrial Electronics*, vol. 58, pp. 4385–4395, 2011.

[12] K. Suratsayadee, J. Himanshu, W. Lee, and C. Kwan, "Wireless health monitoring system for vibration detection of induction Motors," in *Proceedings of the IEEE Industrial and Commercial Power Systems Technical Conference*, pp. 1–6, 2010.

[13] H. Zhang, P. Zanchetta, K. J. Bradley, and C. Gerada, "A low-intrusion load and efficiency evaluation method for in-service motors using vibration tests with an accelerometer," *IEEE Transactions on Industry Applications*, vol. 46, pp. 1341–1349, 2010.

[14] G. Anastasi, M. Conti, and M. A. Di Francesco, "Comprehensive analysis of the MAC unreliability problem in IEEE 802.15.4 wireless sensor networks," *IEEE Transactions on Industrial Informatics*, vol. 7, pp. 52–65, 2011.

[15] L. Hou and N. W. Bergmann, "Novel industrial wireless sensor networks for machine condition monitoring and fault diagnosis," *IEEE Transactions on Instrumentation and Measurement*, vol. 61, pp. 2787–2798, 2012.

[16] V. C. Gungor and G. P. Hancke, "Industrial wireless sensor networks: challenges, design principles, and technical approaches," *IEEE Transactions on Industrial Electronics*, vol. 56, pp. 4258–4265, 2009.

[17] A. Willig, K. Matheus, and A. Wolisz, "Wireless technology in industrial networks," *Proceedings of the IEEE*, vol. 93, pp. 1130–1151, 2005.

[18] J. L. A. Santos, R. C. C. Araujo, A. C. Lima Filho, F. A. Belo, and J. A. G. Lima, "Telemetric system for monitoring and automation of railroad networks," *Transportation Planning and Technology*, vol. 34, pp. 593–603, 2011.

[19] J. Lee, "Performance evaluation of IEEE 802.15.4 for low-rate wireless personal area networks," *IEEE Transactions on Consumer Electronics*, vol. 52, pp. 742–749, 2006.

[20] P. Cheong, K. Chang, Y. Lai, S. Ho, I. Sou, and K. Tam, "A zigbee-based wireless sensor network node for ultraviolet detection of flame," *IEEE Transactions on Industrial Electronics*, vol. 58, pp. 5271–5277, 2011.

[21] Z. Hanzálek and P. Jurčík, "Energy efficient scheduling for cluster-tree wireless sensor networks with time-bounded data flows: application to IEEE 802.15.4/ZigBee," *IEEE Transactions on Industrial Informatics*, vol. 6, no. 3, pp. 438–450, 2010.

[22] C. Caione, D. Brunelli, and L. Benini, "Distributed compressive sampling for lifetime optimization in dense wireless sensor networks," *IEEE Transactions on Industrial Informatics*, vol. 8, no. 1, pp. 30–40, 2012.

[23] J. M. R. Ascariz and L. Boquete, "System for measuring power supply parameters with ZigBee connectivity," in *Proceedings of the IEEE Instrumentation and Measurement Technology (IMTC '07)*, pp. 1–5, May 2007.

[24] A. C. L. Filho, F. A. Belo, J. L. Alves Dos Santos, and E. Gomes Dos Anjos, "Self-powered telemetric torque meter," *Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME*, vol. 133, no. 4, Article ID 045001, 2011.

[25] W. Guo, W. M. Healy, and M. Zhou, "An experimental study of interference impacts on ZigBee-based wireless communication inside buildings," in *Proceedings of the IEEE International Conference on Mechatronics and Automation (ICMA '10)*, pp. 1982–1987, August 2010.

[26] Microchip—MiWi Protocol, 2013, http://www.microchip.com/miwi/.

[27] J. Song, S. Han, A. K. Mok et al., "WirelessHART: applying wireless technology in real-time industrial process control," in *Proceedings of the 14th IEEE Real-Time and Embedded Technology and Applications Symposium (RTAS '08)*, pp. 377–386, April 2008.

[28] S. Petersen and S. Carlsen, "Performance evaluation of WirelessHART for factory automation," in *Proceedings of the IEEE Conference on Emerging Technologies and Factory Automation (ETFA '09)*, pp. 1–9, September 2009.

[29] J. Frey and T. Lennvall, "Wireless sensor networks for automation," in *Networked Embedded Systems*, R. Zurawski, Ed., pp. 27.1–27.43, CRC Press, 2009.

[30] ISA100, Wireless Systems for Automationÿ, 2013, http://www.isa.org/isa100.

[31] A. Flammini, D. Marioli, E. Sisinni, and A. Taroni, "Design and implementation of a wireless fieldbus for plastic machineries," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 3, pp. 747–755, 2009.

[32] F. Salvadori, M. de Campos, P. S. Sausen et al., "Monitoring in industrial systems using wireless sensor network with dynamic power management," *IEEE Transactions on Instrumentation and Measurement*, vol. 58, no. 9, pp. 3104–3111, 2009.

[33] L. Hou and N. W. Bergmann, "Induction motor condition monitoring using industrial wireless sensor networks," in *Proceedings of the 6th International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP '10)*, pp. 49–54, December 2010.

[34] J. Hu, "The application of wireless sensor networks to in-service motor monitoring and energy management," in *Proceedings of the 1st International Conference on Intelligent Networks and Intelligent Systems (ICINIS '08)*, pp. 165–169, November 2008.

[35] H. Jingtao, "In-service motor monitoring and energy management system based on wireless sensor networks," in *Proceedings of the 11th International Conference on Electrical Machines and Systems (ICEMS '08)*, pp. 823–826, October 2008.

[36] C. A. Boano, N. Tsiftes, T. Voigt, J. Brown, and U. Roedig, "The impact of temperature on outdoor industrial sensornet applications," *IEEE Transactions on Industrial Informatics*, vol. 6, no. 3, pp. 451–459, 2010.

[37] L. Lo Bello and E. Toscano, "Coexistence issues of multiple co-located IEEE 802.15.4/ZigBee networks running on adjacent radio channels in industrial environments," *IEEE Transactions on Industrial Informatics*, vol. 5, pp. 157–167, 2009.

[38] L. Tang, K.-C. Wang, Y. Huang, and F. Gu, "Channel characterization and link quality assessment of IEEE 802.15.4-compliant radio for factory environments," *IEEE Transactions on Industrial Informatics*, vol. 3, no. 2, pp. 99–110, 2007.

[39] A. Willig, "Recent and emerging topics in wireless industrial communications: a selection," *IEEE Transactions on Industrial Informatics*, vol. 4, no. 2, pp. 102–122, 2008.

[40] V. C. Gungor, B. Lu, and G. P. Hancke, "Opportunities and challenges of wireless sensor networks in smart grid," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 10, pp. 3557–3564, 2010.

[41] J. Åkerberg, M. Gidlund, and M. Bjorkman, "Future research challenges in wireless sensor and actuator networks targeting industrial automation," in *Proceedings of the 2011 9th IEEE International Conference on Industrial Informatics (INDIN '11)*, pp. 410–415, July 2011.

[42] J. S. Hsu and B. P. Scoggins, "Field test of motor efficiency and load changes through air-gap torque," *IEEE Transactions on Energy Conversion*, vol. 10, pp. 471–477, 1995.

[43] J. S. Hsu, J. D. Kueck, M. Olszewski, D. A. Casada, P. J. Otaduy, and L. M. Tolbert, "Comparison of induction motor field efficiency evaluation methods," *IEEE Transactions on Industry Applications*, vol. 34, no. 1, pp. 117–125, 1998.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 539250, 7 pages
http://dx.doi.org/10.1155/2013/539250

Research Article

A PEM Fuel Cell Diagnostic System Using the Extension Theory

Chin-Tsung Hsieh, Meng-Hui Wang, and Ying-Piao Kuo

Department of Electrical Engineering, National Chin-Yi University of Technology, No. 35, Lane 215, Section 1, Chung-Shan Road, Taiping City, Taichung County 411, Taiwan

Correspondence should be addressed to Meng-Hui Wang; wangmh@ncut.edu.tw

Received 26 September 2012; Revised 31 May 2013; Accepted 2 June 2013

Academic Editor: Tai-hoon Kim

Composed of a single proton exchange membrane fuel cell (PEMFC), a sensor module, and a ZigBee wireless communication module, a fault diagnostic system is proposed in this work to monitor the operation of a fuel cell system. Accordingly, such quantities as cell's output voltage, current, operating temperature, and pressures of gases supplied are monitored. Subsequently, an extension matter-element model is built according to malfunctions of the fuel cell system, which are further categorized into 7 types, each with 12 sorts of characteristics. An extension evaluation method is then directly applied to diagnose such fuel cell system. A human machine interface built under LabVIEW 2009 is incorporated in such a way that a fault(s) can be detected and fixed in a timely manner such that the life cycle of such fuel cell system can be extended.

1. Introduction

As a significant progress as well as a rapid economic growth is made in human society, a tremendous amount of natural resources had been consumed already, leading to issues of immediate concern, in particular the global oil crisis and warming effect. Hence, the developments of green energy technologies are seen more critical than ever before. Among a number of alternative energy sources, hydrogen is treated as one of the most promising candidates, due to the reason that it can be burnt directly to generate heat and can be even applied to a fuel cell as an input to provide electricity through electrochemical reaction. On top of that, a high conversion efficiency up to 40~60% is seen in a fuel cell. Besides, as a consequence of technology improvement and progress made in material science, the power density of a fuel cell has been elevated largely, and fuel cells have been turned into a competitive product in market to a great extent owing to successful cost reduction activities on electrode catalysts and other key components [1].

As a device designed to convert chemical energy to electricity, a fuel cell is mainly composed of three parts, an anode, a layer electrolyte, a cathode. As such, the nature of a fuel cell is subject to the electrolyte contained and chemical mechanism. Two water molecules are formed as the outcome of a chemical reaction between two hydrogen molecules and an oxygen molecule, that is, a pollution free chemical process. Featuring a low pollution level, the development and applications of fuel cell-related technologies have received global attention [2]. However, a fuel cell performance is found as a function of the fuel purity, flow rate, and operating temperature, among other quantities.

Accordingly, aiming to develop a fault diagnostic system for a fuel cell, this work employs extension theory to precisely locate a fault(s). Through a wireless link via a ZigBee module and a GPRS module, a distant monitoring system is reached by way of Ethernet network.

2. Principles and Models of Fuel Cell

As referred previously, a fuel cell is able to convert chemical energy into electrical form, and chemical formulae are represented in (1) and (2). Illustrated in Figure 1 is an electricity generation system of PEMFC, an electrochemical and thermodynamic model [3–5], according to which a potential fault might be located as follows:

$$H_2 \longrightarrow 2H^+ + 2e^-, \tag{1}$$

$$4H^+ + 4e^- + O_2 \longrightarrow 2H_2O \tag{2}$$

The output voltage V provided is expressed as

$$V = E_{\text{thermo}} - V_{\text{act}} - V_{\text{ohmic}} - V_{\text{con}}, \tag{3}$$

where E_{thermo} represents the reversible voltage and V_{ohmic} represents the ohmic voltage drop. The ohmic loss can be minimized by use of a thinner layer of electrolyte and a high conductivity material as follows:

$$E_{\text{thermo}} = 1.229 - 0.85 * 10^{-3}(T - 298.15) + 4.31 * 10^{-5} T * \left[\ln\left(P_{H_2}\right) + 0.5 \ln\left(P_{O_2}\right)\right], \tag{4}$$

$$V_{\text{act}} = -\left[\xi_1 + \xi_2 T + \xi_3 \ln\left(CO_2\right) + \xi_4 T \ln\left(I_{\text{FC}}\right)\right]. \tag{5}$$

In (4), P_{H_2} and P_{O_2}, respectively, represent the pressures of hydrogen and oxygen gases, and T represents the operating temperature of a fuel cell, while in (5) V_{act} denotes the voltage drop across activation of the anode and cathode, ξ_i (i = 1–4) is the cell characteristic coefficient, I_{FC} is the cell current, and CO_2 (atm) is the oxygen concentration as follows:

$$V_{\text{ohmic}} = I_{\text{FC}}\left(R_M + R_C\right), \tag{6}$$

$$R_m = \frac{\sigma M * l}{\text{A}}, \quad \sigma \text{M} = \frac{a}{b}, \quad a = 181.6\left[1 + 0.03\left(\frac{I_{\text{FC}}}{A}\right) + 0.062\left(\frac{T}{303}\right)^2\left(\frac{I_{\text{FC}}}{A}\right)^{2.5}\right], \quad b = \left[\Psi - 0.634 - 3\left(\frac{I_{\text{FC}}}{A}\right)\right]\exp\left[4.18\left(T - \frac{303}{T}\right)\right], \tag{7}$$

$$V_{\text{con}} = -B \ln\left(1 - \frac{J}{J_{\max}}\right), \tag{8}$$

Equation (6), R_C symbolizes the electronic current resistance, and R_M symbolizes the resistance of a PEM, while in (7) l, A and Ψ represent the thickness, the area, and the membrane resistance coefficient, respectively, and in (8) V_{con} denotes concentration loss in the diffusion process, $B(V)$ denotes the operation constant of any type of fuel cell, J denotes the current density, and $J_{\max}$ denotes the maximum current density.

From the above equations, cell's output voltage is found to increase with the operating temperature. A cooling system would be damaged in case the fuel cell is operated beyond rated temperature [6]. A MATLAB simulation, as presented in Figure 2, demonstrates cell's performance with the operating temperature T as a parameter. In the case of $\Psi < 23$, the cell is deemed undermoisturized, not optimized for electricity generation. Faults are detected through 12 characteristics from all the above equations.

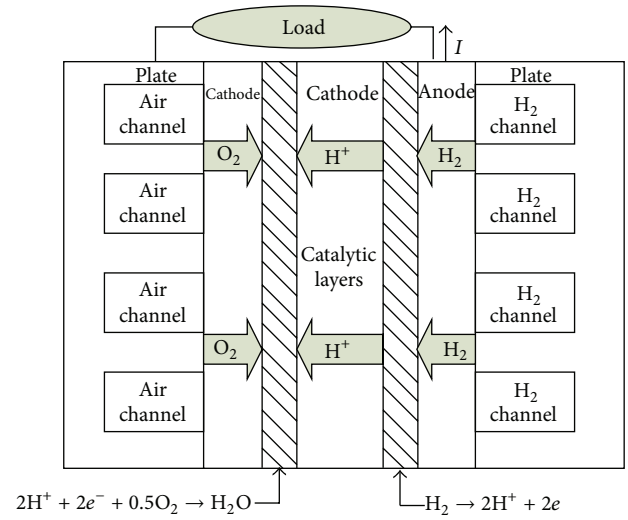


Figure 1: A schematic diagram of a PEMFC power generating system.

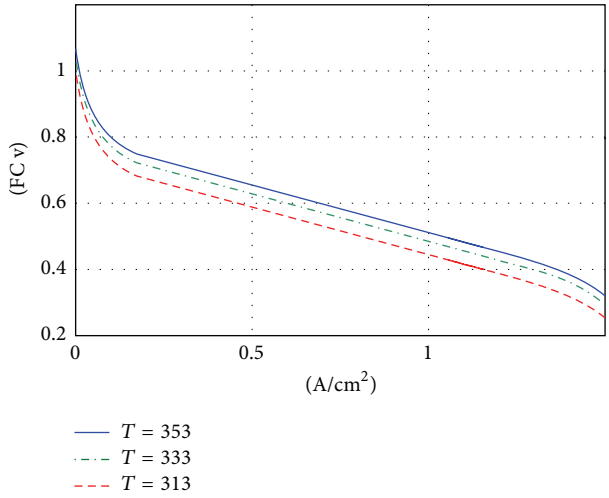


Figure 2: A plot of a single cell's performance under various temperatures.

3. The Fault Diagnostic System Architecture

3.1. The Full Cell System Architecture. The fault diagnostic system proposed in this work is made up of an electricity generation system, a sensor module, and a ZigBee wireless communication module. Sensed data, such as cell's voltage V, current A, the operating temperature T, and the pressure of supplied gas PH, are linked via the ZigBee module to a PC, a PDA, or a smart phone for monitoring purpose with a user friendly interface implemented in LabVIEW 2009. Sketched in Figure 3 is a system configuration, and an entity is pictured in Figure 4.

Widely applied to industrial automation, Modbus, developed by MODICON, is an open and standard communication protocol [7]. ZigBee is a wireless communication module stipulated by IEEE 802.15.4 and ZigBee Alliance in both

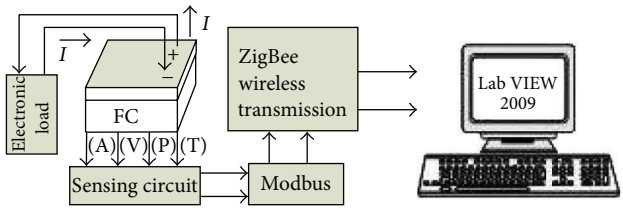


Figure 3: A configuration of a fuel cell monitoring system.

Figure 4: A system entity photo.

the hardware and software aspects, mainly applied to wireless sensor networks, industrial automation, health care, and so forth.

3.2. The Fuel Cell Failure Characteristics. Currently, fuel cells can be roughly categorized into two types: the first of which is operated based on the electrochemical reaction between pure hydrogen and oxygen, while the second is based on the reaction between the pure hydrogen and the oxygen contained in the air. The latter is further classified into two types, namely, self-cooled and pressurized. The one adopted in this work belongs to self-cooled for an advantage of being portable due to the absence of bottled oxygen. However, a major disadvantage accompanied is a short life cycle relative to pressurized. As can be found from a prior work [8], faults may be attributed to a number of factors, for example, overheating as a consequence of a cooling fan malfunction. As many as 12 characteristics are detected with a sensor module in this work. Figure 5 is a configuration of the diagnostic system, where sensor spots are marked on the surface of a gas bottle.

The characteristics acquired are applied to the diagnostic system proposed for the fault identification purpose. As tabulated in Table 1, F_2 signifies an exhaust malfunction, giving rise to a drop in cell's output voltage. Indicating a malfunction in the cooling system, F_3 and F_4 represent an underheated fuel cell system and an overheated one, respectively. F_5 and F_6 denote a malfunction in the gas supply system, while F_7 represents an interrupted wireless network link.

Other than the inlet pressure and the bottle temperature, all the characteristics are applied to (9) for evaluating each changes, where Δt represents the sampling interval as follows:

$$n_i = \frac{[X(i+1) - X(i)]}{\Delta t}. \tag{9}$$

It is found that (5) does not as expected provide a satisfactory identification result.

Instead, the 12 characteristics, as tabulated in Table 2, are applied to (10). As illustrated in Figure 6, $\overline{H_i}$, evaluated as the mean rates of change at 5 instants, that is, $m_i, \ldots, m_{i-4}$, is applied to the proposed approach for fault identification. Consequently, a real time monitoring system can be achieved, according to which the faults in a fuel cell system can be precisely identified in a timely manner as follows:

$$\overline{H_i} = \frac{[m_{i-4} + m_{i-3} + m_{i-2} + m_{i-1} + m_i]}{5}. \tag{10}$$

The fault diagnostic system proposed is built with an interface implemented in LabVIEW 2009. Pictured in Figure 7 is a window of such monitoring system, and each type of malfunction is indicated by individual indicator. Besides, there are two levels of diagnosis contained in such system: the first of which is built for the system level as shown in Figure 8, the second level of the diagnosis system can show the fault location of the fuel cell as shown in Figure 9. As presented in Figure 9, a blockade of the oxygen supply system is indicated in block 5.

4. The Signal Fault Diagnosis Method

Proposed in 1983 by Tsai Wen, the well-developed extension theorem has been successfully applied to a wide range of research fields, for example, artificial intelligence, decision making skill, biomedical engineering, testing technology, and so forth. It extends the binary logic into a continuous and multivalued form. Besides, a correlation function is employed as a way to represent the mature of a thing, that is, the extent that a element belongs to Characteristics of a thing is represented by a real number between $-\infty$ and ∞, a number referred to as the membership grade for such element related to Things characteristic belongs to a collection. After normalization, a membership grade of 1 indicates that element matches completely thing feature, while –1 indicates the exact opposite, and that between –1 and 1 represents an extent somewhere between the previous two cases [9, 10].

4.1. The Extension Matter Element. The term "thing" in everyday life is referred to as "name" for research purpose. A distinctive nature of a thing is characterized by a characteristic, which is quantized as a number referred to as "value." A thing is represented by a set of characteristics, name and the value thereof. As a fundamental unit to describe a thing, a matter element, given a characteristic, a name, and a value, is represented as

$$R = (N, C, V). \tag{11}$$

Due to $V = C(N)$, the relationship between the Value and the Characteristic of a Matter-element, (12) is rewritten as

$$V = (N, C, C(N)). \tag{12}$$

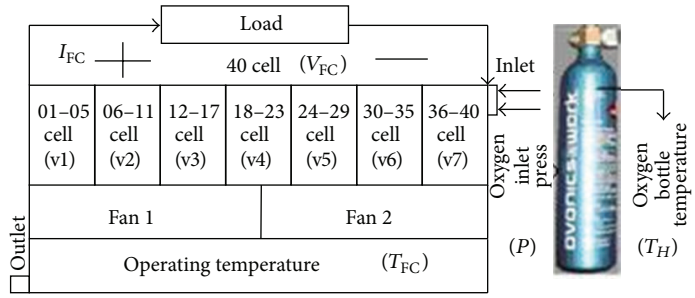


Figure 5: Sensor spots in a fuel cell monitoring system.

Table 1: Fault types in a fuel cell diagnostic system.

F_1	Normal system
F_2	System exhaust valve failure
F_3	System operating temperature lose body heat
F_4	Cooling system failure
F_5	Oxygen holes to plug
F_6	For the hydrogen system failure
F_7	Communications system failure

Table 2: Twelve characteristics in a system.

Feature	Name
C_1	$\overline{V_{FC}}$
C_2	$\overline{I_{FC}}$
C_3	01–05 cell ($\overline{V_1}$)
C_4	06–11 cell ($\overline{V_2}$)
C_5	12–17 cell ($\overline{V_3}$)
C_6	18–23 cell ($\overline{V_4}$)
C_7	24–29 cell ($\overline{V_5}$)
C_8	30–35 cell ($\overline{V_6}$)
C_9	36–40 cell ($\overline{V_7}$)
C_{10}	T_{FC}
C_{11}	T_H
C_{12}	P

In extension theory, $R = (N, C, V)$ can be a multidimensional matter element, as expressed in (13). It contains a characteristic vector $C_M = [C_1, C_2, \ldots, C_n]$, corresponding to a characteristic vector $V_M = [V_1, V_2, \ldots, V_n]$, $R_j = (N, C_j, V_j)$, and $j = 1, 2, \ldots, n$, the submatter element of R. Any matter in daily life can be described in a concrete or an abstract manner, modeled as

$$R = \begin{Bmatrix} N & C_1 & V_1 \\ & \vdots & \vdots \\ & C_n & V_n \end{Bmatrix} = \begin{bmatrix} R_1 \\ \vdots \\ R_n \end{bmatrix}. \quad (13)$$

A matter-element space is portrayed in Figure 10, with the x, y, and z axes representing N, C, and V, respectively.

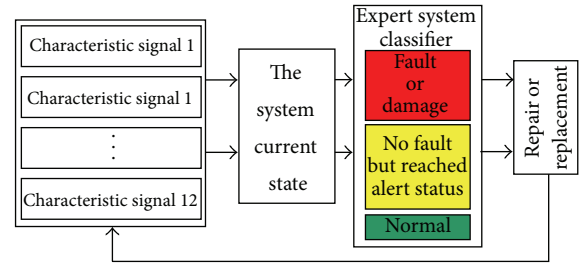


Figure 6: A signal flow graph of a fuel cell diagnostic system.

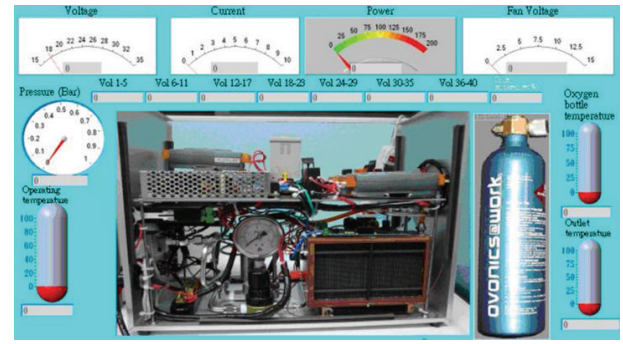


Figure 7: A human machine interface.

4.2. The Extension Evaluation Method. Underlain by an extension set and a correlation function, Extenics is developed as a mathematic tool in an effort to apply such theory to practical applications. The approach procedure is stated as follows.

Step 1. First of all, each fault type is modeled as

$$R_i = \begin{bmatrix} F_i & C_1 & V_{i1} \\ & C_2 & V_{i2} \\ & \vdots & \vdots \\ & C_n & V_{in} \end{bmatrix}, \quad i = 1, 2, \ldots, 7;\ n = 1, 2, \ldots, 12, \quad (14)$$

where C_i represents each set of characteristics in level i, V_{0ji} represents the distribution range covered by characteristic i in level j, and a_{0ji} and b_{0ji} represents the maximum and the minimum of such characteristic set in the corresponding level. Accordingly, a thing R is decomposed into a set of

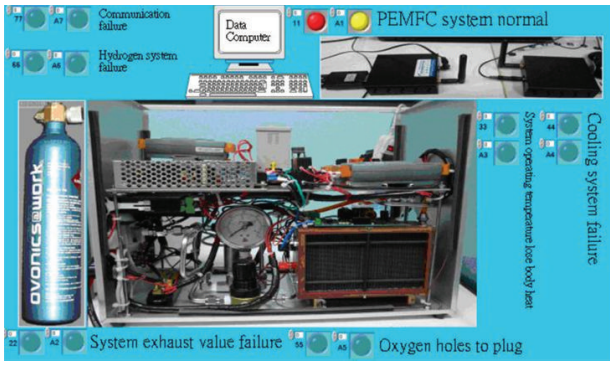


FIGURE 8: The first level of a diagnostic system.

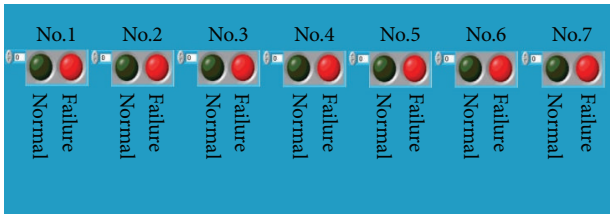


FIGURE 9: The second level of a diagnostic system.

section field RP, and j is decomposed into sets of classical field R_{0j}, expressed as

$$R_{0j} = \left(N_{0j}, C_i, X_{0ji}\right) = \begin{bmatrix} N_{0j} & C_1 & V_{0j1} = \left\langle a_{0j1}, b_{0j1}\right\rangle \\ & C_2 & V_{0j2} = \left\langle a_{0j2}, b_{0j2}\right\rangle \\ & \vdots & \vdots \\ & C_{12} & V_{0j12} = \left\langle a_{0j12}, b_{0j12}\right\rangle \end{bmatrix}. \quad (15)$$

Step 2. Referred to as the evaluated matter element, a set of characteristics pertains to a matter element R, represented as

$$R = \left(q, C_i, x_i\right) = \begin{bmatrix} q & C_1 & v_1 \\ & C_2 & v_2 \\ & \vdots & \vdots \\ & C_{12} & v_{12} \end{bmatrix}, \quad (16)$$

where q denotes a set of characteristic values and x_i the value of C_i in q, that is, the specific information available in a evaluated things. In other words, a single thing R can be characterized by multiple sets of characteristic values q.

Step 3. Each fault correlation function is evaluated as

$$K_{ij}\left(v_{tj}\right) = \begin{cases} -\dfrac{0.5\rho\left(v_{tj}, V_{ij}\right)}{\left|V_{ij}\right|}, & \text{if } v_{tj} \in V_{ij}, \\ \dfrac{\rho\left(v_{tj}, V_{ij}\right)}{\left[\rho\left(v_{tj}, V'_{pj}\right) - \rho\left(v_{tj}, V_{ij}\right)\right]}, & \text{if } v_{tj} \notin V_{ij}, \end{cases} \quad (17)$$

$$i = 1, 2, \ldots, 7; \quad j = 1, 2, \ldots, 12.$$

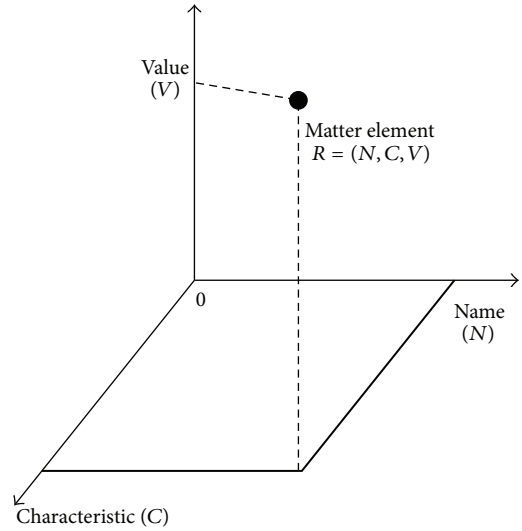


FIGURE 10: A matter-element space.

Step 4. The weighting factor λ_i, ranging from 0 to 1, is defined as a measure of the significance of C_i to R, and the total sum of λ_i is identically unity, as expressed in

$$\lambda_i = \sum_{j=1}^{12} W_{ij} K_{ij}; \quad i = 1, 2, \ldots, 7. \quad (18)$$

Step 5. Finally, all the weighted correlation functions $\lambda_i\ k_j(x_i)$ are summed up, that is, $k_j(q)$. As given in (19), the maximum value of λ_i is selected as the evaluation results in type j as follows:

$$\lambda_{\max} = \max_{1 \le i \le 7} \left\{\lambda_i\right\}. \quad (19)$$

Step 6. Normalization is performed in (20) in order that the fault diagnosis value falls within the interval $\langle 1, -2\rangle$ as intended each time as follows:

$$\lambda'_i = \frac{3\lambda - \lambda_{\min} - 2\lambda_{\max}}{\lambda_{\max} - \lambda_{\min}}, \quad i = 1, 2, \ldots, 7, \quad (20)$$

where

$$\lambda_{\max} = \max_{1 \le i \le 7} \left\{\lambda_i\right\}; \qquad \lambda_{\min} = \min_{1 \le i \le 7} \left\{\lambda_i\right\}. \quad (21)$$

Step 7. A fault is identified as belonging to type if in case $\lambda f' = 1$. The identification is made totally according to the correlation thereof, due to the presumption that a high level of correlation implies a high possibility that a corresponding type of fault occurs.

Step 8. In case all the parts in a fuel cell system have been diagnosed once, then the diagnostic procedure comes to an end otherwise skip back to Step 2 for another run.

The idea of extension evaluation method is that the experiment data accumulated are classified into a certain number of level collections, to which respective ranges of real

Table 3: Performance and requirement comparison among various approaches.

Name	Times of learning	Learning recognition rate	Test the recognition rate
Identification method in this paper	0	99.37%	98.75%
K_{means}	0	69.02%	40%
Neural network (12-16-7)	1000	82.81%	66.25%
Neural network (12-14-7)	1000	89.37%	78.75%
Neural network (12-12-7)	1000	88.75%	70%
Neural network (12-10-7)	1000	85.81%	63.75%

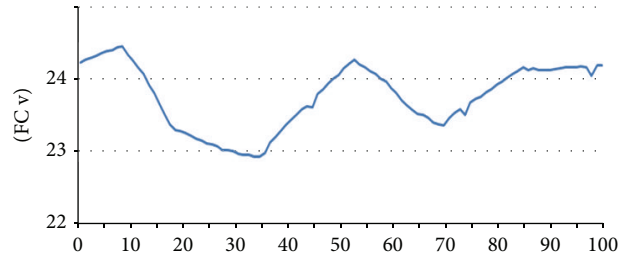


Figure 11: A voltage response to a malfunction in a temperature control system.

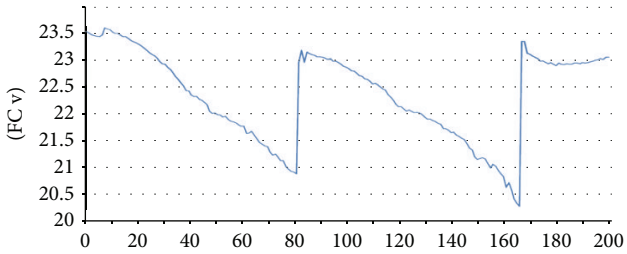


Figure 12: A voltage response to a malfunction in exhaust valve.

numbers are assigned. Subsequently, the pending assessment of the data is applied to each range of assigned numbers successively for the evaluation of corresponding membership grade. A higher membership grade indicates a higher level of linkage between the pending assessment of the data and such level collections and vice versa.

5. Experimental Results and Discussion

5.1. The Detector Signal. Underlain by a mathematic model built for a fuel cell and the characteristics thereof, out of which the characteristics of objects is extracted, the type of a fault(s) is identified through extension theory and grey system theory. A malfunction of a fuel cell system is reflected by a drop in the output voltage provided. Not taking proper measures in time may result in a permanent damage to the cell system. In this work, a signal data base is constructed in Excel for futuristic system diagnosis, according to which output voltage signals are plotted against time. Presented in Figures 11 and 12 are the curves indicating faults in the temperature control system (F_3 and F_4) and the exhaust value (F_2), respectively. As many as 100 and 200 data records are made with a sampling interval of 5 seconds, a tunable quantity

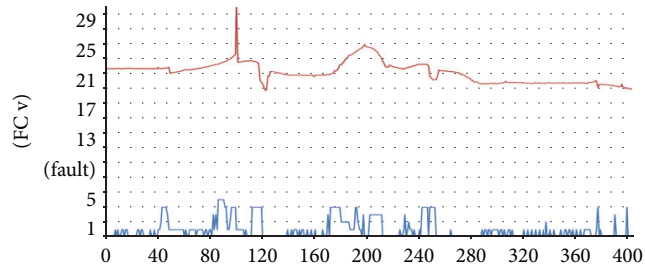


Figure 13: A plot of identification result against time.

through the human-machine interface. The trend similarity between such two curves in Figures 11 and 12 is seen, meaning that there is no way to precisely identify the fault types in the absence of a systematic diagnostic approach. For this sake, this work is proposed as efficient means to identify faults and take required actions in a timely manner against any sort of potential damage to the cell system.

5.2. The Diagnostic Signals. As demonstrated in Figure 13, the curve in red indicates cell's output voltage versus time, while in blue indicates the corresponding diagnosis result. The y coordinates represent the identification results, namely, fault types $F_1 = 1$, $F_2 = 2$, $F_3 = 3$, $F_4 = 4$, $F_5 = 5$, and $F_6 = 6$, at discrete time instants.

Up to 80% of data records are treated as the training samples, and the rest are as the test samples. As tabulated in Table 3, the approach proposed, not requiring a training process, is found superior to k-means classifier in terms of identification rate and superior to a variety of neural network, necessitating a training process, as well.

6. Conclusions

Presented in this work is a fault diagnostic system for a fuel cell, an easy to implement system made up of multiple Modbus modules. On top of that, a user friendly human machine interface is constructed for easy monitoring of the cell system operation. Over others, the approach proposed, not requiring a training process, acquires advantage of high recognition rate, meaning that a fault(s) in an early stage can be identified in a timely manner in order that measures can be taken to extend the life span of the cell system. Integrated with a ZigBee wireless communication module, this system

can be in the future applied to a distant monitoring system, such as an alter system for a fuel cell-powered vehicle.

Acknowledgment

The authors feel deeply indebted to the National Science Council, Taiwan. for all the resources gained under Grant no. NSC 99-2221-E-167-031-MY2.

References

[1] J. Larminie and A. Dicks, *Fuel Cell Systems Explained*, John Wiley & Sons, New York, NY, USA, 2003.

[2] J. M. Corrêa, F. A. Farret, L. N. Canha, and M. G. Simões, "An electrochemical-based fuel-cell model suitable for electrical engineering automation approach," *IEEE Transactions on Industrial Electronics*, vol. 51, no. 5, pp. 1103–1112, 2004.

[3] A. Rowe and X. Li, "Mathematical modeling of proton exchange membrane fuel cells," *Journal of Power Sources*, vol. 102, no. 1-2, pp. 82–96, 2001.

[4] W. Zhen, *The characteristics of simulation studies of proton exchange membrane fuel cell systems [M.S. thesis]*, Energy and Power Engineering Institute of Shandong University, 2007.

[5] N. Yousfi-Steiner, Ph. Moçotéguy, D. Candusso, D. Hissel, A. Hernandez, and A. Aslanides, "A review on PEM voltage degradation associated with water management: impacts, influent factors and characterization," *Journal of Power Sources*, vol. 183, no. 1, pp. 260–274, 2008.

[6] S. H. Ge and B. L. Yi, "A mathematical model for PEMFC in different flow modes," *Journal of Power Sources*, vol. 124, pp. 1–11, 2003.

[7] L. A. M. Riascos, M. G. Simões, and P. E. Miyagi, "A Bayesian network fault diagnostic system for proton exchange membrane fuel cells," *Journal of Power Sources*, vol. 165, no. 1, pp. 267–278, 2007.

[8] M.-H. Wang, K.-H. Chao, G. J. Huang, and H.-H. Tsai, "Application of extension theory to fault diagnosis of automotive engine," *ICIC Express Letters*, vol. 5, no. 4B, pp. 1293–1299, 2011.

[9] M.-H. Wang, Y.-K. Chung, and W.-T. Sung, "The fault diagnosis of analog circuits based on extension theory," in *Emerging Intelligent Computing Technology and Applications*, vol. 5754 of *Lecture Notes in Computer Science*, pp. 735–744, 2009.

[10] K.-H. Chao, S.-H. Ho, and M.-H. Wang, "Modeling and fault diagnosis of a photovoltaic system," *Electric Power Systems Research*, vol. 78, no. 1, pp. 97–105, 2008.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 294967, 6 pages
http://dx.doi.org/10.1155/2013/294967

Research Article

Correction of Nonlinear Frequency Sweep in Frequency-Modulated Continuous-Wave Laser Range Sensor

Soo-Yong Jung,[1] Seong Ro Lee,[2] and Chang-Soo Park[1]

[1] *School of Information and Communications, Gwangju Institute of Science and Technology, 261 Cheomdan-gwagiro, Buk-gu, Gwangju 500-712, Republic of Korea*
[2] *Department of Information and Electronics Engineering, Mokpo National University, Yeongsan-ro, Cheonggye-myeon, Jeollanam-do, Muan-gun 1666, Republic of Korea*

Correspondence should be addressed to Chang-Soo Park; csp@gist.ac.kr

Received 1 February 2013; Revised 10 June 2013; Accepted 11 June 2013

Academic Editor: Tai-hoon Kim

We propose a linearization method for reducing the effect of nonlinear frequency sweep in a frequency-modulated continuous-wave (FMCW) based laser range sensor. In FMCW laser range sensors, nonlinear frequency sweep can severely degrade the measurement accuracy because it gives the system ambiguity when determining the target range. In general, voltage controlled oscillators (VCO) which are used for frequency modulation show nonlinear frequency sweep property even though the input voltage signal is a linear ramp signal. To solve this problem, we adopt an additional fixed delay structure to extract the nonlinearity and compensate it. The proposed linearization method has been worked out through the numerical process and the simulation, and this method effectively eliminates the nonlinear frequency sweep problem.

1. Introduction

Laser range sensors are remote distance sensing devices with typical applications, such as solid-target detections, 3D vision, localization, and robotics [1–3]. To determine the distance, projecting an optical signal onto an object and processing the reflected or scattered signal are performed in laser range sensors. Conventionally, pulsed time-of-flight (TOF), phase-shift measurement, and frequency-modulated continuouswave (FMCW) are considered as major techniques for laser range sensors [4]. Pulsed TOF range sensor determines the distance by measuring the round trip time of the optical pulse signal. This method provides high signal-to-noise ratio (SNR) and shows good performance; however high cost and large size are the drawbacks [5, 6]. Phase-shift range sensor determines the distance using phase difference between reference and reflected signals. Using this method, simple and low cost laser range sensor can be realized. However, intermediate frequency drift and influence of the crosstalk degrade the performance of the phase-shift technique [7, 8].

To avoid these problems, we focused on the FMCW laser range sensor. In an FMCW laser range sensor, a sinusoidal signal with a constant rate of frequency change is transmitted, and this signal is reflected by a target. The frequency difference between the transmitted and the reflected signals, called the beat frequency, contains the distance information. As long as the frequency sweep is linear, the beat frequency is focused at a single frequency, and the target distance can be easily extracted. However, in practice, the linear frequency sweep profile is not easy to obtain. If there is nonlinearity in the frequency sweep, the beat frequency is not focused at a single frequency, and it is difficult to determine the exact range [9–11]. There are several techniques for linearization of the frequency sweep, and most of them focused on the linearization of the voltage controlled oscillator (VCO) frequency sweep because it is the crucial nonlinearity source of the system. The techniques are mainly of two types: one is open-loop correction [12, 13] and the other one is closed-loop correction [14–16]. The open-loop correction method modifies the VCO tuning voltage properly to get a linear frequency sweep using a look-up table. However, since the frequency is

controlled only by an open-loop system, inevitable frequency drifts, due to temperature, environmental conditions, aging, and so forth, cannot be compensated for. The closed-loop correction method adopts a phase locked loop (PLL) circuitry to linearize the frequency sweep. However, because PLLs are dynamic systems, any changes in the system input will cause the output not to follow immediately but to exhibit some transient behavior. Such errors in the output frequency will have a negative influence on the FMCW measurement result.

To solve these problems, we propose a method for correction of frequency-sweep nonlinearity in a signal processor instead of linearization of the VCO frequency sweep. For linearization, an additional fixed delay structure is adopted, and the frequency-sweep nonlinearity is extracted and used for compensation. We validate our linearization method with numerical analysis and computer simulation.

2. Nonlinearity Correction Method

In FMCW technique, a transmitter produces an FMCW signal, that is, a sinusoidal signal with a constant rate of frequency change. This signal is backscattered by a target and goes back to the emitter. Beat frequency which is frequency difference between transmitted signal and backscattered signal contains the distance information.

In intensity-modulated FMCW laser range sensor, a laser diode (LD) transmits an optical signal with direct modulation, and a VCO is adopted for frequency modulation. The output signal frequency of the VCO is controlled by input voltage. If the input voltage of VCO is a ramp signal, an optical signal with a constant rate of frequency increase is generated. This optical signal is backscattered by a target and goes back to a photo detector (PD) which can transfer optical signal to electrical signal. Using a mixer and a low pass filter (LPF), a signal with beat frequency is extracted. Then the distance of the target can be expressed as

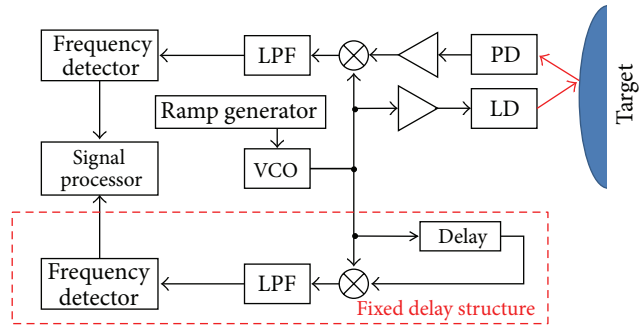
$$d = \frac{\tau_{\mathrm{RTT}} \cdot c}{2} = \frac{f_{\mathrm{BEAT}} \cdot t_{\mathrm{RAMP}} \cdot c}{2 \cdot \Delta f}, \tag{1}$$

where τ_{RTT} is the round trip time of the optical signal, c is the speed of light, f_{BEAT} is the beat frequency, t_{RAMP} is period of the ramp signal, and Δf is frequency sweep range of the VCO. In third term of (1), t_{RAMP}, c, and Δf are known parameters, and the beat frequency, f_{BEAT}, is determined by the target distance. If the frequency sweep is linear, the beat frequency is constant as long as the target range is not changed. However, in practice an exact ramp-like frequency profile is not easy to obtain. If there is nonlinearity in the frequency sweep, it makes problem because the beat frequency is not constant, and it is difficult to find exact range.

To solve the nonlinear sweep problem, we propose equipping the FMCW laser range sensor with an additional fixed delay structure, as depicted in Figure 1.

As shown in Figure 1, a fixed delay structure is added for linearization. When a voltage ramp signal is given to the VCO, the angular frequency of the VCO output signal can be expressed as

$$\omega(t) = \omega_0 + \gamma(t) \cdot t, \tag{2}$$



LD: laser diode
VCO: voltage controlled oscillator
PD: photo detector
LPF: low pass filter

Figure 1: Block diagram of the proposed FMCW laser range sensor.

where ω_0 is the initial angular frequency and $\gamma(t)$ means the tuning rate to be compensated due to nonlinearity. For simplicity, assuming $\Delta\gamma(t) \ll t$, the following equation shows the phase of the VCO output signal:

$$\varphi(t) = \int_0^t \omega(t)\, dt = \varphi_0 + \omega_0 t + \frac{1}{2}\gamma(t) \cdot t^2, \tag{3}$$

where φ_0 is the initial phase. Then, the VCO output signal can be written as

$$E_{\mathrm{VCO}}(t) = K_{\mathrm{VCO}} \cos(\varphi(t)), \tag{4}$$

where K_{VCO} is the amplitude of the VCO output signal. This signal is modulated to the laser diode and transmitted to the target. Practically, the LD output optical power is a nonlinear function of the input electrical signal. Although we can minimize the nonlinearity by modulating the input signal in the linear region of LD, the residual components can remain. The output from the LD is shown in (5) for a third-order nonlinearity, where m is the optical modulation index, and A_2 and A_3 are device dependent nonlinearity coefficients [17]:

$$\begin{aligned} P_T(t) &= P_0 \left[1 + mK_{\mathrm{VCO}} \cos(\varphi(t)) \right. \\ &\quad + A_2 m^2 K_{\mathrm{VCO}}^2 \cos^2(\varphi(t)) \\ &\quad \left. + A_3 m^3 K_{\mathrm{VCO}}^3 \cos^3(\varphi(t))\right] \\ &= P_0 \left[1 + \frac{B_2}{2} + \left\{B_1 + \frac{3B_3}{4}\right\} \cos(\varphi(t)) \right. \\ &\quad \left. + \frac{B_2 \cos(2\varphi(t))}{2} + \frac{B_3 \cos(3\varphi(t))}{4}\right], \end{aligned} \tag{5}$$

where $B_1 = mK_{\mathrm{VCO}}$, $B_2 = A_2 m^2 K_{\mathrm{VCO}}^2$, and $B_3 = A_3 m^3 K_{\mathrm{VCO}}^3$. Then, the reflected signal detected by PD is

expressed as [18]

$$\begin{aligned} E_{\mathrm{PD}}(t) &= K\left|RP_R(t)\right|^2 \\ &= K\left|R\frac{n+1}{2\pi d^2}\rho A_R \sin^2(\mathrm{FOV})\, P_T\left(t-\tau_{\mathrm{RTT}}\right)\right|^2 \\ &= C_0 + C_1\cos\left(\varphi\left(t-\tau_{\mathrm{RTT}}\right)\right) + C_2\cos\left(2\varphi\left(t-\tau_{\mathrm{RTT}}\right)\right) \\ &\quad + C_3\cos\left(3\varphi\left(t-\tau_{\mathrm{RTT}}\right)\right) + C_4\cos\left(4\varphi\left(t-\tau_{\mathrm{RTT}}\right)\right) \\ &\quad + C_5\cos\left(5\varphi\left(t-\tau_{\mathrm{RTT}}\right)\right) + C_6\cos\left(6\varphi\left(t-\tau_{\mathrm{RTT}}\right)\right), \end{aligned} \tag{6}$$

where $P_R(t)$ is the received optical signal, R is responsivity of the PD, K is constant of proportionality, n is a mode number, d is the target range, ρ is a reflection coefficient, A_R is the photosensitive area of the PD, FOV is a field of view, and τ_{RTT} is the round trip time of the transmitted signal. The VCO output signal and the reflected signal are multiplied through a mixer, and a signal with the beat frequency is obtained after LPF, expressed as

$$E_{\mathrm{BEAT}}(t) = \frac{K_{\mathrm{VCO}}C_1}{2}\cos\left(\varphi(t)-\varphi\left(t-\tau_{\mathrm{RTT}}\right)\right). \tag{7}$$

A frequency detector is used for detecting the frequency of the signal after LPF. A reciprocal frequency counter can be used as the frequency detector. Because the reciprocal frequency counter measures the period for one cycle of the waveform, it can support high resolution and very fast readings. Also, a Schmitt trigger circuit can be used at input stage of the frequency detector so that noise does not cause spurious edges. Because frequency is derivative of phase, the angular frequency to be detected is expressed as

$$\begin{aligned} \omega_{\mathrm{BEAT}}(t) &= \frac{d}{dt}\left[\varphi(t)-\varphi\left(t-\tau_{\mathrm{RTT}}\right)\right] \\ &= \tau_{\mathrm{RTT}}\cdot\gamma(t) + \tau_{\mathrm{RTT}}\frac{d\gamma(t)}{dt}t - \frac{{\tau_{\mathrm{RTT}}}^2}{2}\frac{d\gamma(t)}{dt}. \end{aligned} \tag{8}$$

As we can see in (8), if the tuning rate $\gamma(t)$ is constant, then the angular frequency difference, $\omega_{\mathrm{BEAT}}(t)$, is equal to the product of the round trip time and the tuning rate. Because $\omega_{\mathrm{BEAT}}(t)$ can be detected by the frequency detector and the tuning rate is a known value, it is easy to find the round trip time, and we can extract the range of thetarget. However, usually, the frequency sweep of a VCO is not perfectly linear; that is, the tuning rate is not constant. Then, the beat frequency cannot give the exact round trip time.

To solve this problem, we added a fixed delay structure for obtaining the tuning rate. The frequency detector can detect the frequency difference between the VCO output and delayed output signal, expressed as

$$\begin{aligned} \omega_{\mathrm{AUX}}(t) &= \frac{d}{dt}\left[\varphi(t)-\varphi\left(t-\tau_D\right)\right] \\ &= \tau_D\cdot\gamma(t) + \tau_D\frac{d\gamma(t)}{dt}t - \frac{\tau_D^2}{2}\frac{d\gamma(t)}{dt}, \end{aligned} \tag{9}$$

where τ_D is the time delay in the fixed delay structure and is much smaller than the frequency sweep time of the ramp signal. In (9), $\omega_{\mathrm{AUX}}(t)$ can be detected by a frequency detector with the time delay. Thus, using the ordinary differential equation, the tuning rate, $\gamma(t)$, can be calculated as

$$\begin{aligned} \gamma(t) &= \exp\left(-\int\frac{2}{2t-\tau_D}dt\right) \\ &\quad\cdot\left[\int\frac{2}{2t-\tau_D}\frac{\omega_{\mathrm{AUX}}(t)}{\tau_D}\exp\left(\int\frac{2}{2t-\tau_D}dt\right)dt\right] \\ &= \exp\left(-\ln\left|2t-\tau_D\right|\right) \\ &\quad\cdot\left[\int\frac{2}{2t-\tau_D}\frac{\omega_{\mathrm{AUX}}(t)}{\tau_D}\exp\left(\ln\left|2t-\tau_D\right|\right)dt\right] \\ &= \frac{1}{\left|2t-\tau_D\right|}\cdot\left[\frac{2}{\tau_D}\int\frac{\left|2t-\tau_D\right|}{2t-\tau_D}\omega_{\mathrm{AUX}}(t)\,dt\right]. \end{aligned} \tag{10}$$

Because the time delay is much smaller than frequency sweep time of the ramp signal as mentioned above, t is larger than $\tau_D/2$ in (10). Accordingly, the tuning rate can be

$$\gamma(t) = \frac{1}{2t-\tau_D}\cdot\left[\frac{2}{\tau_D}\int_{\tau_D/2}^{t}\omega_{\mathrm{AUX}}(t)\,dt\right]. \tag{11}$$

The derivative of the tuning rate can be obtained from (9) and expressed as

$$\frac{d\gamma(t)}{dt} = \frac{2}{2t-\tau_D}\left(\frac{\omega_{\mathrm{AUX}}(t)}{\tau_D}-\gamma(t)\right). \tag{12}$$

Using (8) and (12), we can obtain a quadratic equation for τ_{RTT}, which is expressed as

$$\begin{aligned} &\frac{1}{2}\frac{2}{2t-\tau_D}\left(\frac{\omega_{\mathrm{AUX}}(t)}{\tau_D}-\gamma(t)\right)\tau_{\mathrm{RTT}}^2 \\ &\quad-\left\{\gamma(t)+\frac{2t}{2t-\tau_D}\left(\frac{\omega_{\mathrm{AUX}}(t)}{\tau_D}-\gamma(t)\right)\right\}\tau_{\mathrm{RTT}} \\ &\quad+\omega_{\mathrm{BEAT}}(t)=0. \end{aligned} \tag{13}$$

The solutions of (13) are expressed as

$$\begin{aligned} \tau_{\mathrm{RTT}} &= \frac{\left(2t/\left(2t-\tau_D\right)\right)\left(\omega_{\mathrm{AUX}}(t)/\tau_D-\tau_D\gamma(t)/2t\right)}{\left(2/\left(2t-\tau_D\right)\right)\left(\omega_{\mathrm{AUX}}(t)/\tau_D-\gamma(t)\right)} \\ &\quad\pm\frac{\sqrt{\left\{\left(2t/\left(2t-\tau_D\right)\right)\left(\omega_{\mathrm{AUX}}(t)/\tau_D-\tau_D\gamma(t)/2t\right)\right\}^2-\left(\left(4\cdot\omega_{\mathrm{BEAT}}(t)\right)/\left(2t-\tau_D\right)\right)\left(\omega_{\mathrm{AUX}}(t)/\tau_D-\gamma(t)\right)}}{\left(2/\left(2t-\tau_D\right)\right)\left(\omega_{\mathrm{AUX}}(t)/\tau_D-\gamma(t)\right)}. \end{aligned} \tag{14}$$

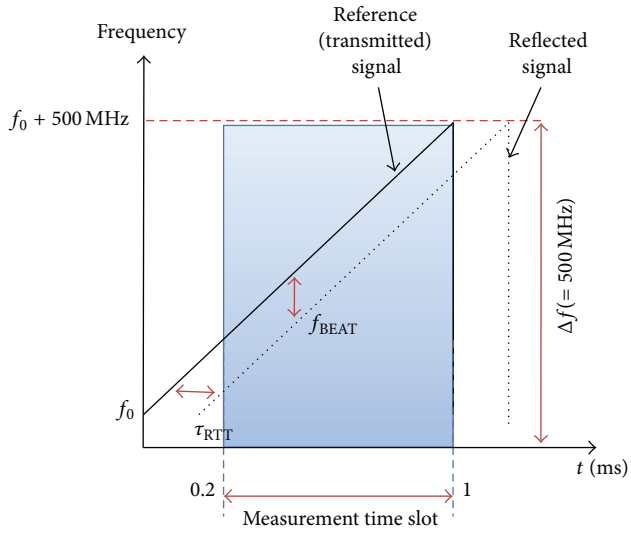


FIGURE 2: Simulation condition.

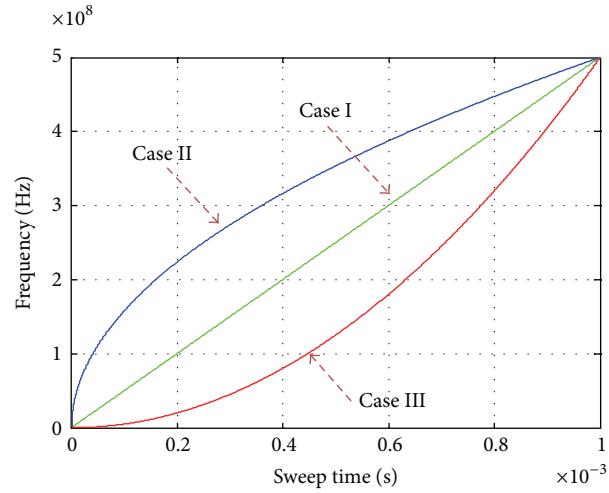


FIGURE 3: Three kinds of frequency sweep patterns for each tuning rate.

There are two solutions, and the solution in the detectable range is selected. The obtained round trip time is constant over the measurement time, lying between the maximum round trip time and the frequency sweep time. In (14), t can be any value of the measurement time. The target range can be obtained using a simple relation between distance and time expressed as

$$d = \frac{\tau_{RTT} \cdot c}{2}. \tag{15}$$

3. Results and Discussion

We evaluate the proposed method using a computer simulation. The frequency sweep range and sweep time of a ramp signal are 500 MHz and 1 ms, respectively, and the measurement time slot is 0.2 to 1 ms. Figure 2 shows the simulation condition.

To evaluate the proposed method, we modeled a nonlinear frequency sweep and applied it to the FMCW laser range sensor. Figure 3 shows three kinds of frequency sweep patterns for each tuning rate.

We modeled one ideal linear and two kinds of nonlinear sweep patterns. Table 1 lists the tuning models.

When the target range is 30 m, the beat frequencies and obtained target range for each case are shown in Figures 4(a) and 4(b), respectively. As shown in Figure 4(a), when the frequency sweep is linear, the beat frequency is focused at a single frequency, 100 kHz, and the target range, 30 m, can be easily obtained. However, when the frequency sweep is nonlinear, it is difficult to extract the exact beat frequency, and the measurement accuracy is degraded. For case II, the obtained target range in the measurement time can vary between 15 and 34 m, and, for case III, the obtained target range can vary between 12 and 58 m, as depicted in Figure 4(b).

TABLE 1: Tuning models.

Case	Tuning rate
I	$\gamma(t) = 500 \times 10^9$ (linear)
II	$\gamma(t) = 500 \times 10^{7.5} t^{-1/2}$
III	$\gamma(t) = 500 \times 10^{12} t$

Two kinds of nonlinear frequency sweep patterns are applied in the proposed method. Figures 5(a) and 5(b) show the simulation results of the obtained target range plotted against the measurement time after linearization for case II and case III, respectively. As we can see that, in both results, the target range is almost constant over the measurement time and can be obtained successfully.

4. Conclusions

We have proposed a linearization method using an additional fixed delay structure in intensity-modulated FMCW laser range sensor. For correction of the nonlinear frequency sweep problem, a fixed structure was adopted to extract the tuning rate, and the target range was calculated using the obtained tuning rate. We modeled three kinds of frequency sweep patterns and applied them to the FMCW laser range sensor. When the frequency sweep was linear, the beat frequency was focused at a single frequency, and the target range was easily extracted. On the other hand, when the frequency sweep was nonlinear, there were multiple frequency components that could not extract the exact target range. From the proposed linearization method, the simulation results clearly showed that the proposed method effectively eliminates the nonlinear frequency sweep problem and improves the measurement accuracy.

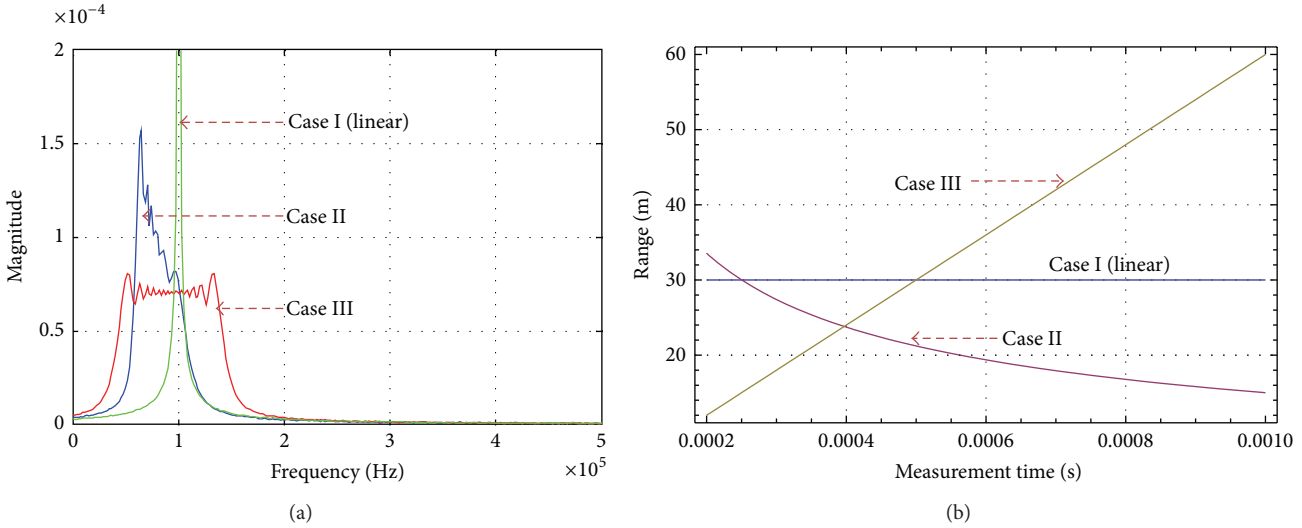


FIGURE 4: (a) Beat frequencies and (b) obtained target range for the tuning rate models.

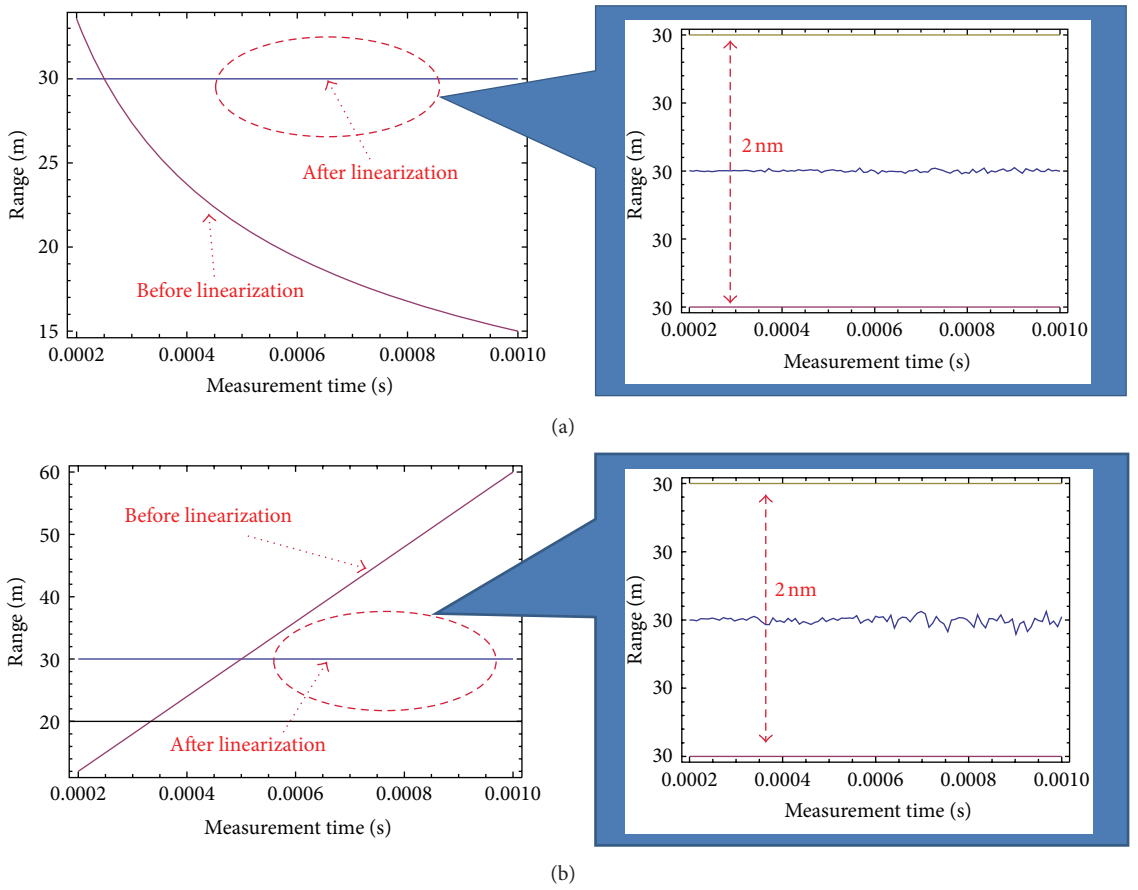


FIGURE 5: Simulation results after application of the proposed linearization method for the nonlinear frequency sweep of (a) $\gamma(t) = 500 \times 10^{7.5} t^{-1/2}$ and (b) $\gamma(t) = 500 \times 10^{12}\, t$.

Acknowledgments

This work was supported by Priority Research Centers Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education, Science and Technology (2009-0093828), and by the National Research Foundation of Korea (NRF) Grant funded by the Korean government (MEST) (no. 2011-0017081).

References

[1] D. C. Carmer and L. M. Peterson, "Laser radar in robotics," *Proceedings of the IEEE*, vol. 84, no. 2, pp. 299–320, 1996.

[2] F. Blais, "Review of 20 years of range sensor development," *Journal of Electronic Imaging*, vol. 13, no. 1, pp. 231–243, 2004.

[3] P. A. Forrester and K. F. Hulme, "Laser rangefinders," *Optical and Quantum Electronics*, vol. 13, no. 4, pp. 259–293, 1981.

[4] M.-C. Amann, T. Bosch, M. Lescure, R. Myllylä, and M. Rioux, "Laser ranging: a critical review of usual techniques for distance measurement," *Optical Engineering*, vol. 40, no. 1, pp. 10–19, 2001.

[5] S. M. Nejad and S. Olyaee, "Unified pulsed laser range finder and velocimeter using ultra-Fast time-to-digital converter," *Iranian Journal of Electrical & Electronic Engineering*, vol. 5, no. 2, pp. 112–121, 2009.

[6] P. Palojärvi, K. Määttä, and J. Kostamovaara, "Pulsed time-of-flight laser radar module with millimeter-level accuracy using full custom receiver and TDC ASICs," *IEEE Transactions on Instrumentation and Measurement*, vol. 51, no. 5, pp. 1102–1108, 2002.

[7] T. Bosch and M. Lescure, "Crosstalk analysis of 1 m to 10 m laser phase-shift range finder," *IEEE Transactions on Instrumentation and Measurement*, vol. 46, no. 6, pp. 1224–1228, 1997.

[8] G. Perchet and M. Lescure, "Magnification of phase shift for a laser rangefinder: intrinsic resolution and improved circuit," *Journal of Optics*, vol. 29, no. 3, pp. 229–235, 1998.

[9] B. L. Stann, W. C. Ruff, and Z. G. Sztankay, "Intensity-modulated diode laser radar using frequency-modulation/continuous-wave ranging techniques," *Optical Engineering*, vol. 35, no. 11, pp. 3270–3278, 1996.

[10] T.-J. Ahn, J. Y. Lee, and D. Y. Kim, "Suppression of nonlinear frequency sweep in an optical frequency-domain reflectometer by use of Hilbert transformation," *Applied Optics*, vol. 44, no. 35, pp. 7630–7634, 2005.

[11] C. J. Karlsson and F. A. A. Olsson, "Linearization of the frequency sweep of a frequency-modulated continuous-wave semiconductor laser radar and the resulting ranging performance," *Applied Optics*, vol. 38, no. 15, pp. 3376–3386, 1999.

[12] D. A. Williams, "A highly linearized mm-wave voltage controlled oscillator for FMCW radar applications," *IEE Colloquium on Solid State Components for Radar*, pp. 601–615, 1988.

[13] H.-G. Park, B. Kim, and Y. S. Kim, "VCO nonlinearity correction scheme for a wideband FM-CW radar," *Microwave and Optical Technology Letters*, vol. 25, no. 4, pp. 266–269, 2000.

[14] T. Musch, I. Rolfes, and B. Schiek, "A highly linear frequency ramp generator based on a fractional divider phase-locked-loop," *IEEE Transactions on Instrumentation and Measurement*, vol. 48, no. 2, pp. 634–637, 1999.

[15] M. Pichler, A. Stelzer, P. Gulden, C. Seisenberger, and M. Vossiek, "Frequency-sweep linearization for FMCW sensors with high measurement rate," in *Proceedings of the IEEE MTT-S International Microwave Symposium*, pp. 1693–1696, June 2005.

[16] S. Scheiblhofer, S. Schuster, and A. Stelzer, "Signal model and linearization for nonlinear chirps in FMCW radar SAW-ID tag request," *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 4, pp. 1477–1483, 2006.

[17] X. N. Fernando and A. B. Sesay, "Adaptive asymmetric linearization of radio over fiber links for wireless access," *IEEE Transactions on Vehicular Technology*, vol. 51, no. 6, pp. 1576–1586, 2002.

[18] F. R. Gfeller and U. Bapst, "Wireless in-house data communication via diffuse infrared radiation," *Proceedings of the IEEE*, vol. 67, no. 11, pp. 1474–1486, 1979.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 249598, 11 pages
http://dx.doi.org/10.1155/2013/249598

Research Article

Continuous Monitoring of Water Quality Using Portable and Low-Cost Approaches

Gurkan Tuna,[1] Orhan Arkoc,[2] and Kayhan Gulez[3]

[1] *Computer Programming Department, Edirne Technical Sciences Vocational School, Trakya University, 22020 Edirne, Turkey*
[2] *School of Technical Sciences, Kirklareli University, 39100 Kirklareli, Turkey*
[3] *Control and Automation Engineering Department, Electrical-Electronics Faculty, Yildiz Technical University, 34220 Istanbul, Turkey*

Correspondence should be addressed to Gurkan Tuna; gurkantuna@trakya.edu.tr

Received 9 January 2013; Revised 13 May 2013; Accepted 13 May 2013

Academic Editor: Tai-hoon Kim

Water quality refers to the physical, chemical, and biological characteristics of water, and it is a measure of the condition of water relative to any human need or purpose. A particular problem with measuring the condition of water quality at drinking water reservoirs is the requirement of collecting a large number of samples. To handle this problem, we focus on the practical use of two different portable and low-cost approaches for continuous monitoring of water quality: miniboats loaded with sondes with probes and wireless sensor network- (WSN-) based monitoring system. These approaches bring several advantages over traditional monitoring systems in terms of cost, portability, and applicability. Our simulation studies show that these systems can be used to monitor water quality at drinking water reservoirs such as dams and holding ponds. Field tests to prove the effectiveness of the proposed systems are in progress.

1. Introduction

Water quality is one of the main factors to control health and the state of diseases in people and animals. Both natural processes such as soil erosion and weathering, and anthropogenic inputs such as industrial and municipal wastewater discharge largely determine surface water quality in a region [1]. Lakes and rivers are the main sources of drinking water, fishery, irrigation, and energy production, which considerably depend on water quality. Therefore, water quality of lakes and rivers should be kept at a certain level. Strategies are being developed to deal with the eutrophication of lakes and rivers with excess nutrients and the pollution of surface water with toxic chemicals around the world [2].

Sustainable use of water resources requires surface water assessment monitoring programs in addition to decision making and management tools. The major reference to guide efforts to attain a sustainable aquatic environment in Europe is the Water Framework Directive (WFD) [3]. The WFD consists of guidelines defining the categories of quality and the required components and parameters. Dissolved oxygen (DO), pH, electrical conductivity (EC), temperature, turbidity, and nitrate are the main parameters to determine the water quality as stated by WFD [3], US EPA [4], and Turkish Regulations [5].

A particular problem in the case of surface water quality monitoring is the complexity associated with collecting a large number of samples. Monitoring of water quality can be performed using different approaches such as fixed continuous monitoring systems, sampling with portable devices, and postanalysis. Fixed monitoring systems require high initial start-up cost but they provide real-time data. On the other hand, sampling with portable devices does not provide real-time water quality data. In this paper, we propose two different approaches for autonomous monitoring of water quality. In the first system, the boats receive trajectories set by the operator before a mission and follow the trajectories to visit predefined sampling points as shown in Figure 1. During the mission, the sondes, water quality monitoring instruments, on the boats analyze water quality and log results. In the second system, a group of portable water quality probes mounted on buoys at fixed positions regularly analyze water quality and send the data to the control center as shown in Figure 2.

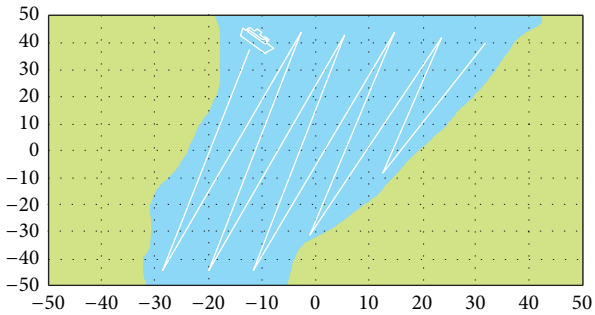


Figure 1: A simple figure to illustrate the first scenario. In the figure, white lines indicate the trajectory followed by the autonomous boat.

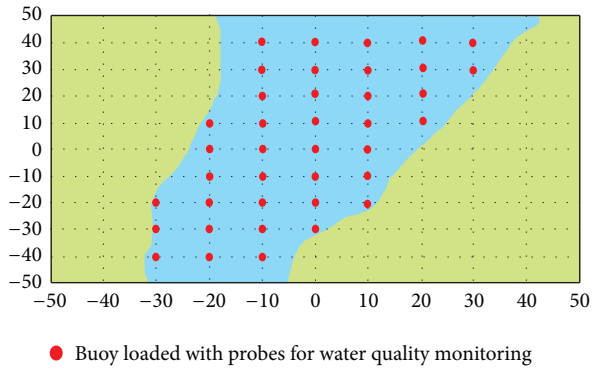


Figure 2: A simple figure to illustrate the second scenario. In this figure, red points represent buoys which carry portable water quality monitoring probes.

The proposed systems have three main advantages over fixed monitoring systems as the following.

(i) The proposed systems do not require a high initial cost and maintenance costs.

(ii) Both systems are portable and can be used at different sites.

(iii) WSN-based water quality monitoring system provides online data to utility providers. On the other hand, monitoring of water quality by using an autonomous boat provides offline data for on-site analyses. This system can be converted to provide online data by adding a two-channel long range radio modem in order to transmit boat position and heading, and water quality data at the same time.

The remainder of this paper is organized as follows. Section 2 introduces different water quality monitoring systems in the literature. Section 3 introduces the use of autonomous miniboats for continuous monitoring of water quality. Section 4 introduces the details of a WSN-based water quality monitoring system. Performance evaluations of the proposed systems are reported in Section 5. A discussion on the proposed approaches is given in Section 6. Finally, the paper is concluded in Section 7.

2. Related Work

Since water quality monitoring with portable devices provides advantages like a low initial start-up cost and low periodical maintenance costs, they are being used widely. Multiprobe water quality sondes with sensors for EC, temperature, pH, and luminescent DO were used in a research at Nam Co Lake, China [6]. A case study in which water quality sondes were used to measure DO during the water quality analysis of Toenepi Stream, New Zealand is reported [7]. Lapen et al. conducted a survey with water quality sondes to examine nutrient quality and bacteria in tile drainage and shallow ground water [8]. In another study, the details of three automatic monitoring stations which record the responses of lakes and reservoirs to changes in the weather are explained [9].

A detailed review of advancements in water quality sensors, telemetry, and computing technologies in addition to real-time remote monitoring applications is stated in [10]. Water quality monitoring networks are essential to characterize and manage water quality accurately and are important in water basin decision making. An important issue for water quality monitoring networks, the design of sampling locations, is explained in [11]. In addition to remote monitoring, another emerging technique for monitoring applications is the use of autonomous unmanned vehicles. In [12], a prototype autonomous underwater vehicle for rapid environmental assessment is discussed. In [13], an unmanned boat for mapping the water quality of shallow mire pools where aquatic weeds flourish is proposed. Though this study is promising in some aspects, it was conducted during daylight hours since it required an operator to row the boat. During daylight hours, sunlight and temperature change. Hence, pH and DO levels change due to photosynthesis. Therefore, this operation should be conducted using a full autonomous boat at night.

Water quality plays an important role not only for humans but also for animals. Water quality monitoring and forecasting is crucial in modern intensive fish farming management. A wireless system for water quality monitoring in fish culture is explained in [14]. Proper management of animal waste treatment lagoons requires regular sludge surveys of the lagoons. In [15], a study in which sludge measurement was conducted by using a GPS enabled sonar equipped airboat is discussed.

In the recent years, many applications have been proposed for WSNs. One of these is precision agriculture, where WSNs are used in several processes including management of water resources for irrigation, understanding the changes in the crops, estimating fertilizer requirements, and predicting crop performance [16]. In [17], the development of a WSN for the inline monitoring of the content of nitrates in a real scenario, River Turia, Spain, is presented. In this real-world application, ion selective electrode transducers and solar panels for energy harvesting were used. In [18], a WSN-based distributed system for chemical analysis of water was proposed. In this study, nitrate, ammonium, and chloride were analyzed at 14 different locations.

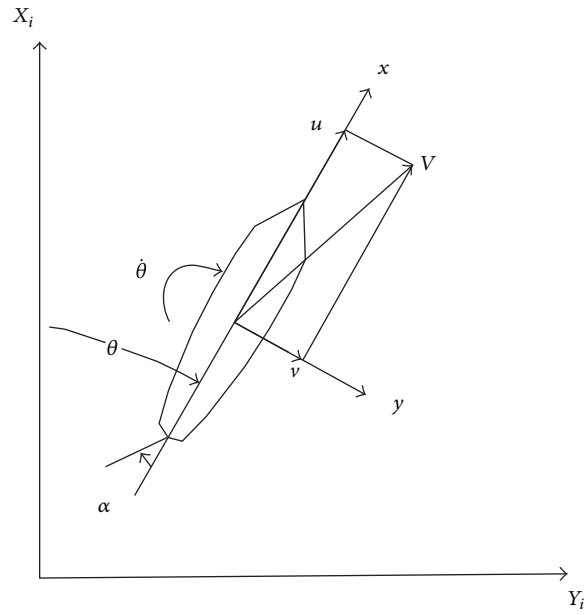


Figure 3: Boat motion description.

In distributed water quality monitoring applications, in addition to communication infrastructures, the development of sensors plays an important role. In [19], in addition to the hardware designs of pH, redox and turbidity sensors, the software design of the proposed approach, the hardware design of a solar panel, and the hardware design of a wireless sensor node are explained. A detailed review of solid state sensors for online monitoring of water quality parameters including pH, DO, EC, turbidity, dissolved metal ions, and dissolved organic carbon on the high spatial resolution is presented in [20].

Different from the approaches in the literature, fully portable water quality monitoring systems are proposed in this study. Autonomous boat-aided water quality monitoring system works fully autonomously without the guidance of an operator. A single system owned by a utility provider can be used at several water reservoirs. WSN-based water quality monitoring system proposed in this study is a portable system and consists of portable monitoring nodes mounted on buoys. Using anchors, buoys are positioned at fixed locations at a water reservoir and form a WSN-based water quality monitoring system. This system can easily be moved to another reservoir if required.

3. The Use of Miniboats to Collect Water Samples and Real-Time Analysis

In this section, the details of the proposed system are explained. Firstly, the design and sailing algorithm of the autonomous boat are given. Then, the proposed navigation system used by the boat is explained. Finally, sensors used to analyze water quality are described as well as the details of logging and analyzing.

3.1. Design and Sailing Algorithm of the Autonomous Miniboat. For autonomous sailing of the boat, a top-down planning system which consists of three subsystems has been designed as follows.

(i) Sensing subsystem: this subsystem is responsible for sensing the environment in order to build a concrete world model. Position, heading, and speed of the boat are measured by the sensors located on the boat. An Inertial Measurement Unit (IMU) provides information about velocity and acceleration. With the aid of a Global Positioning System (GPS) receiver, the IMU determines the position of the boat. An ultrasonic wind sensor provides information about wind speed and direction. Due to the nature of the implementation areas, the AIS system, a common system in commercial ships to ensure that collisions are avoided, has not been installed.

(ii) Trajectory planning subsystem: this subsystem is responsible for generating a plan to reach the goal point.

(iii) Executing subsystem: this subsystem executes required actions of the plan.

The down-down planning system uses an algorithm based on a well-known A* search in x, y, and heading of the boat and takes the nonholonomic properties of the boat, wind speed, and wind direction into account. The paths given by the algorithm minimize sailing duration even in case the boat encounters some static and dynamic obstacles. The subsystems take care of specific requirements of the application. The sensing subsystem has to give information regarding wind speed and wind direction to the other subsystems which calculate the heading. This subsystem is also responsible for detecting moving obstacles. The trajectory planning subsystem has to find a path which minimizes the time to sail towards the target [21]. The executing subsystem has to be robust for long-term implementations. A trajectory planning cycle is composed of a global and a local planning cycle. The global planning cycle generates a rough map from current position to the target. And then, a detailed path is generated by the local planning cycle.

Mathematical model of the boat dynamics and disturbances are necessary to obtain a model for the steering of the boat [22, 23]. The boat in this study is modeled as a rigid body with six degrees of freedom (DOF) corresponding to translations along the three axes and its rotations about these axes as shown in Figure 3. By neglecting the couplings between the motion in the horizontal plane and the roll, pitch, and heave motions, the equations have been simplified. In order to obtain this model two reference frames have been defined as follows.

(i) Inertial reference frame: this frame is fixed to the Earth. The x_i axis points towards the North, the y_i axis towards the East, and the z_i axis towards the centre of the Earth. The origin is located on the mean water surface at an appropriate location.

(ii) Body reference frame: this frame is fixed to the hull.

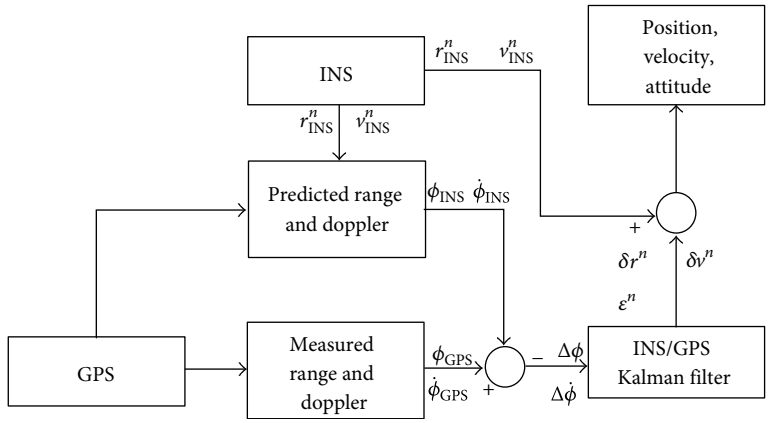


Figure 4: INS-based GPS-aided navigation system—tight integration approach.

The position and heading of the boat can be described as

$$\vec{p} = \begin{pmatrix} x \\ y \\ \theta \end{pmatrix}. \tag{1}$$

The projections of the speed V are called the surge velocity u and the sway velocity v. The heading angle and the turning angle are denoted θ and $\dot{\theta}$. As given in [24], the motion of the boat can be expressed using the following equations:

$$\begin{aligned} X &= m \cdot \left(\dot{u} - v \cdot \dot{\theta} - x_G \cdot \ddot{\theta}\right), \\ Y &= m \cdot \left(\dot{v} - u \cdot \dot{\theta} - x_G \cdot \ddot{\theta}\right), \\ N &= I_Z \cdot \dot{\theta} + m \cdot x_G \cdot \left(\dot{v} + u \cdot \dot{\theta}\right), \end{aligned} \tag{2}$$

where X and Y are the components of the hydrodynamic forces and disturbances, N is the z-component of the moments, m is the mass of the boat, and I_Z is the moment of inertia. By setting the centre line of the boat to the origin, the centre of gravity can be neglected [24].

To quantify the hydrodynamic forces, each part of the boat needs to be analyzed. Though the prototype boat is not a sail, wind effects, the results of the actual flow of air, should be investigated in order to analyze the forces generated by the boat. In addition to the wind effects, rudder forces should be taken into consideration. Due to the difference in the pressure of the water over the lower and upper surfaces of the lifting rudder foil, a lifting action occurs. Also, damping forces, based on the resistance of the keel and hull, exist. The motion models in this study do not consider disturbances. However, boats are influenced by waves, currents, and wind. These disturbances can be modeled by forces and moments. Since possible application areas of the prototype system are all types of drinking reservoirs such as holding ponds and dams, the effects of these forces are neglected to simplify the design of the prototype system and it is left as future work.

3.2. INS-Based GPS-Aided Navigation System. Since the distance of remote control systems is limited and autonomous operation is preferred to teleoperated mode, an INS-based GPS-aided navigation system shown in Figure 4 was designed to enable the boat follow a predetermined trajectory. While the reference trajectory is provided by the integrated IMU, the GPS receiver acts as the updating system. The reason behind this approach is that the frequency of INS measurements is higher than that of GPS measurements. In this study, the geodetic reference for GPS is the World Geodetic System 1984 (WGS84). The Cartesian coordinate frame of reference used in GPS is Earth-Centered Earth-Fixed (ECEF). Three-dimensional coordinates are used by ECEF to describe the location of a GPS user or satellite.

There are two common strategies for the integration of INS data and GPS measurements. The loose integration strategy consists of a GPS filter and an INS filter and these filters operate separately. This strategy operates if one of the filters fail and the processing time required for this strategy is less than tight integration strategies [25, 26]. The tight integration strategy consists of a single filter as shown in Figure 4. Since there is a single filter in this strategy, the state vector increases in size, and this results in longer processing time [25]. On the other hand, this strategy has two main advantages over the loose integration strategy. First, information among states is shared statistically in this strategy. Second, the filtering of the GPS measurements is improved [26].

Both integration approaches can be implemented in two different ways: *closed loop* and *open loop*. In the closed loop implementation, all the error estimates plus misalignment error are used to correct the mechanization parameters which are used during determining the position and velocity provided by the INS. Therefore, small errors are propagated by the mechanization equations. On the other hand, in the open loop implementation, misalignment errors are sent back to the mechanization equations. Therefore, the errors tend to grow. This type of implementation is only suitable for high end INS.

The INS/GPS filter in Figure 4 is the combination of the INS and GPS filters of the loose integration strategy. Since the INS and GPS filters share the same position and velocity states, the INS/GPS filter is basically the INS filter augmented with the double difference ambiguities. In most systems, due to the difference between the locations of the

GPS receiver and the INS, the position and velocity of the INS is different from the one of the GPS receiver. To solve this problem, a vector from the centre of the INS to the GPS antenna is needed. As a result of the time differences of the INS data and GPS measurements, a mechanism to make sure that the position and velocity provided by the INS coincide with the GPS measurements is needed. A common method to handle this problem is linear interpolation. Linear interpolation method is suitable for low dynamic vehicles such as cars, boats, and sails. In case of high dynamic vehicles, Lagrange interpolation is a common method.

The main steps of the tight integration strategy which was adopted from [25, 27] are as follows.

(1) By using the mechanization equations, raw INS measurements are processed in order to determine the position and velocity provided by the INS.

(2) By using the position and velocity obtained in Step 1 and raw GPS ephemeris information, pseudoranges and Doppler measurements are predicted.

(3) The predicted pseudoranges and Doppler measurements obtained in Step 2 are fed into an INS/GPS Kalman filter. In order to determine the error estimates of the position and velocity in addition to misalignment error, the filter takes the difference between the raw GPS pseudorange and Doppler measurements and the pseudoranges and Doppler measurements from Step 2.

(4) Finally, the error estimates obtained in Step 3 are used to update the position and velocity obtained in Step 1 in order to get a full state vector.

In the proposed system, the INS/GPS error states are defined by the following equation:

$$\delta \dot{x} = F\delta x + Gw. \tag{3}$$

To simplify the design of the INS/GPS Kalman filter, the position of the boat is expressed in terms of Cartesian coordinate on East-North-Up (ENU) instead of geodetic coordinate used by the GPS receiver. Then, the position errors along ENU can be written as

$$\begin{bmatrix} \delta r_{\mathrm{E}} \\ \delta r_{\mathrm{N}} \\ \delta r_{\mathrm{U}} \end{bmatrix} = \begin{bmatrix} \delta\lambda \left(R_N + h\right) \cos\varphi \\ \delta\varphi \left(R_M + h\right) \\ \delta h \end{bmatrix}, \tag{4}$$

where r_{E}, r_{N}, and r_{U} represent the Cartesian coordinates on ENU and δr_{E}, δr_{N}, and δr_{U} represent the position errors along ENU. φ, λ, and h represent geodetic coordinates.

The error states are selected as

$$\begin{aligned} &\delta x \\ &= [\, \delta r_{\mathrm{E}} \;\; \delta r_{\mathrm{N}} \;\; \delta r_{\mathrm{U}} \;\; \delta v_{\mathrm{E}} \;\; \delta v_{\mathrm{N}} \;\; \delta v_{\mathrm{U}} \;\; \delta p \;\; \delta r \;\; \delta A \;\; gb_x \;\; gb_y \;\; gb_z \\ &\quad ab_x \;\; ab_y \;\; ab_z \;\; gsf_x \;\; gsf_y \;\; gsf_z \;\; asf_x \;\; asf_y \;\; asf_z \;\; \delta t \;\; \delta \dot{t} \,]^T, \end{aligned} \tag{5}$$

where r is the position vector, v is the velocity vector, p is the pitch angle, A is the heading angle, g is the gravity vector, a is the acceleration vector, b is the accelerometer bias, f is the specific force vector, s is the scale factor, and t is time variable.

The velocity vector along ENU is given as

$$\left[V^{\mathrm{E}} \;\; V^{\mathrm{N}} \;\; V^{\mathrm{U}}\right]^T. \tag{6}$$

The accelerometer measurements on the body frame are given as

$$\left[f_x \;\; f_y \;\; f_z\right]^T. \tag{7}$$

The specific force on the local level frame is

$$\left[f^{\mathrm{E}} \;\; f^{\mathrm{N}} \;\; f^{\mathrm{U}}\right]^T. \tag{8}$$

The stabilization gyro measurements on the body frame is given as

$$\left[w_x \;\; w_y \;\; w_z\right]^T. \tag{9}$$

Detailed information on deriving the Kalman filter equations and adapting the equations for INS systems and GPS receivers can be found in [28, 29].

3.3. Probe Used for Water Analyzing and Logging. Different types of probes are used for water analyses. In this study, we prefer a Sonde with the following probes for water quality analyses. The Sonde logs the data for offline evaluations.

CTD. This probe measures conductivity, temperature, and depth. Maximum permissible limit for conductivity is 2.5 mS/cm according to WFD and maximum allowable temperature for inland water resources is 25°C according to Turkish Regulations, which are in the measurement range of the probe on the prototype Sonde. The ranges for the probe are between 0 to 100 mS/cm for conductivity, −5 to +50°C for temperature, and 0 to 200 m for depth. These values are automatically calculated from conductivity according to algorithms found in Standard Methods for the Examination of Water and Wastewater [30].

DO. Maximum permissible limit for DO is 5 mg/L according to WFD and is 8 mg/L according to Turkish Regulations, which are in the measurement range of the sensor planned to be used. The probe on the prototype Sonde uses the methods approved by both ASTM D888-12 Standard Test Methods for DO in Water [31] and the US EPA and can detect DO between 0 and 50 mg/L. The accuracy of the probe is 0 to 20 mg/L ± 2% of reading or 0.2 mg/L.

pH. pH of the water according to both WFD and Turkish Regulations should be between 6.5 and 8.5 units. The probe on the prototype Sonde measures pH in units from 0 to 14 with an accuracy of ±0.2 unit.

Nitrate. Recommended maximum permissible limit for Nitrate is 50 mg/L according to WFD, 10 mg/L according to US EPA, and 20 mg/L according to Turkish Regulations. The probe on the prototype Sonde measures between 0 and 200 mg/L-N with an accuracy of ±10% of reading.

Table 1: Specifications of the water quality monitoring boards.

Item	DO board	EC board	pH board
Range	0–20 Mg/L	—	0–14 (Na^+ error at >12.3 pH)
Accuracy	+/−0.1	+/−5 μs	+/−0.20 pH
Probe body material	Epoxy and Noryl	Epoxy	Epoxy
Max probe temperature	50°C	70°C	70°C
Max probe PSI	690 kPa (100 PSI)	1379 kPa (200 PSI)	1379 kPa (200 PSI)
Baud rate (default)	38400 bps	38400 bps	38400 bps

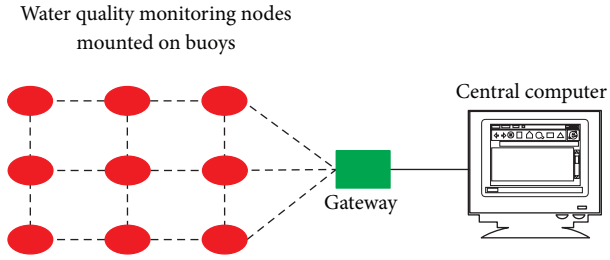


Figure 5: WSN-based water quality monitoring system.

Turbidity. Turbidity has no actual limit according to US EPA and WFD regulations. Turbidity can be measured with the probe on the prototype Sonde between 0 and 1,000 NTU units with an accuracy of ±2% of reading.

All the probes on the prototype Sonde are certified by EPA Environmental Technology Verification Program [32].

4. WSN-Based Water Quality Monitoring

Wireless sensor nodes contain a microcontroller, a storage unit, A/D converters, a radio module, a battery, and one or more sensors for measuring different environmental parameters including temperature, light, acceleration, and angular velocity. Wireless sensor nodes convert data frames to radio messages and send these frames to the gateway. All radio units within a given range establish a wireless sensor network (WSN). In WSNs, nodes automatically establish and maintain connectivity by using mesh-networking protocols. The sensor gateway and its associated middleware connect the WSN with the enterprise computing environment.

In this study, DO, EC, pH development kits of Atlas Scientific [33] have been preferred and integrated with waspmote sensor node development platforms [34]. Waspmote is an open source wireless sensor platform which offers low power consumption modes to allow the sensor nodes to be autonomous and battery-powered up to five years depending on the duty cycle and the radio used. At the moment, a battery-operated custom designed board with an 802.15.4-based wireless network interface has also being developed for water quality monitoring. Figure 5 illustrates the implementation of a WSN-based water quality monitoring system. In this system, portable water quality monitoring nodes monitor water quality and provide wireless connection. A gateway provides connection between the WSN and a central computer which is located at the control center of a utility provider. The central computer provides WSN data repository and is used for analyses. Table 1 lists the specifications of the water quality monitoring boards.

Table 2: Part of the dataset used in the simulation study.

X	Y	ID	EC	Temp (°C)	Nitrate
523710	4620530	1	7	17.7	4
523640	4620530	2	7	17.6	3.8
523640	4620570	3	7	17.8	4
523680	4620570	4	7.8	17.9	3.7
523610	4620650	5	7	17.9	3.6
523550	4620660	6	7.6	18	3.7
523560	4620720	7	8	17.8	3.8
523500	4620800	8	8.2	17.8	3.9
523430	4620890	9	8.1	17.9	3.5
523520	4620900	10	8.4	17.9	4
523620	4620900	11	8.3	18	4
523670	4620950	12	8.7	18	3
523550	4620950	13	8.3	18	3.7

5. Performance Evaluations

In this section, two sets of performance evaluations and our future research directions are reported. First set of performance evaluations was conducted to show the effectiveness of using an autonomous boat for water quality monitoring. Second set of performance evaluations was conducted to evaluate lifetime expectations of portable water quality monitoring nodes. Finally, planned future research directions are reported.

5.1. Simulation Studies on the Use of an Autonomous Boat for Water Quality Monitoring. In order to prove the effectiveness of the proposed INS-GPS integrated navigation system, a simulation environment was developed in MATLAB. INS and GPS data were integrated in after mission mode. It was assumed that the inputs to the simulator were supplied from a GPS receiver and an IMU. In this respect, the GPS receiver provides observation data as binary files and Receiver Independent Exchange Format (RINEX) version 2.10 files, navigation data as binary files and RINEX version 2.10 files, and IMU data as binary files and text files. While observation data consists of the raw code, phase and Doppler measurements, navigation data consists of ephemeris parameters. We imported previously collected RINEX version 2.10 files into MATLAB and edited them accordingly in order to simulate the scenario shown in Figure 1. Figure 6 shows the locations to be visited by the simulated boat for water quality analyses. We conducted twenty sets of simulations in which the boat followed random paths to visit previously defined sampling

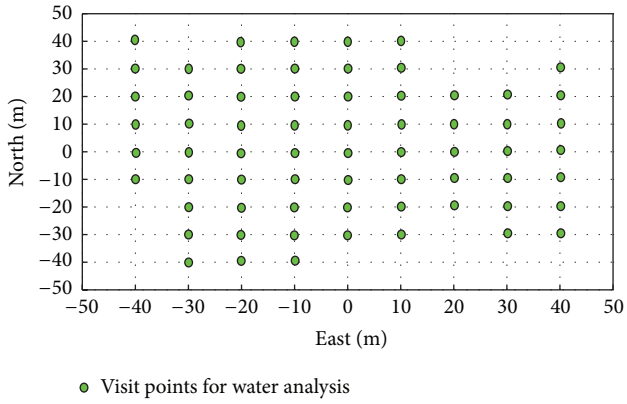


Figure 6: Visit points of the simulated boat for water analyses.

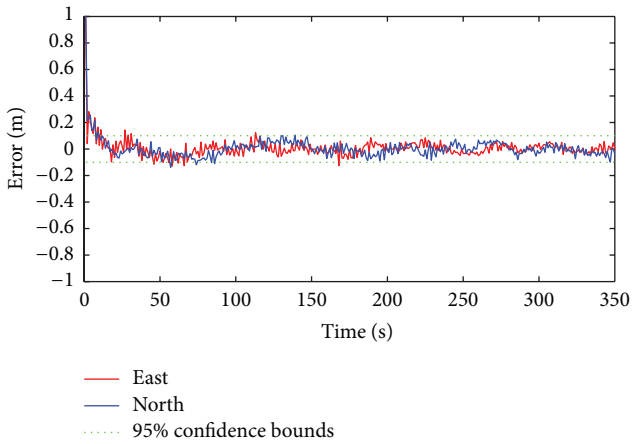


Figure 7: Trajectory errors of the simulated boat from the East and the North (in meters).

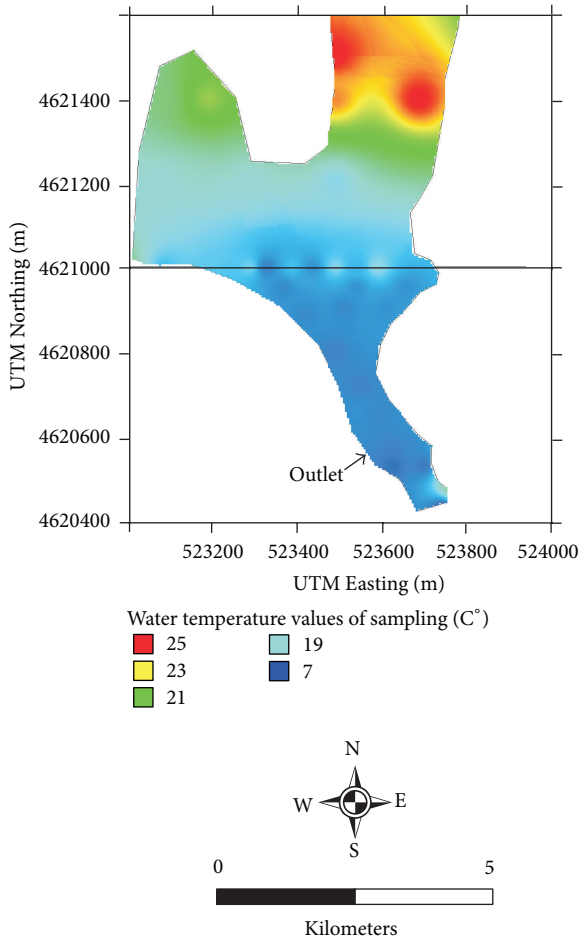


Figure 8: Water temperature values of the sample dataset.

points, green circles, shown in Figure 6. Positional errors (x, y, and heading) of the boat in each simulation were logged in a CSV file and then the plot shown in Figure 7 was obtained. Figure 7 shows the average of the trajectory errors of the simulated boat from the East and the North. The results prove that the boat follows its predefined trajectory successfully. Considering the results of this simulation study and the accuracy of the Sonde, we can conclude that water quality maps of drinking water resources can be obtained by using the proposed system.

In order to show the use of autonomous boats for water quality monitoring applications, a set of Geographic Information System- (GIS-) based simulation studies was conducted by using a previously collected dataset consisting of sampling locations, EC, temperature, and nitrate values. In Table 2, IDs are used to identify sampling locations. This table lists a part of the dataset used in this simulation study. The results of these simulation studies are shown in Figures 8, 9, and 10.

5.2. Simulation Studies on the Lifetime Evaluations of Portable Water Quality Monitoring Nodes. As it is well known, lifetime expectations of a wireless node with internal or external batteries depend on several parameters and facts including battery capacity, transmission power, transmission frequency, duty period, node density, and several node-related parameters. Considering the specific requirements of the proposed approach, measurement rate and transmission power are the main parameters which affect the lifetime of a monitoring operation. To analyze these parameters, we developed a MATLAB-based simulation environment using the source codes of MATSNL [35] and analyzed the effects of measurement rate and transmission power on node lifetime.

In the first case study, we set the specifications of the simulated mote according to the specifications of telosB [36] motes. We set the transmission power of the motes to 0 dBm and −25 dBm and varied the measurement rate from one minute to one day. The resulting plot is shown in Figure 11. As in Figure 11, when sensor measurement rate is increased, node lifetime drops significantly.

In the second case study, we again set the specifications of the simulated mote according to the specifications of telosB motes and computed the power breakdown of the average power consumed by a node for transmitting, receiving, sensing, preprocessing, and sleeping. We varied the duty cycle, the ratio between active period, and the full active/dormant period of a sensor node, between 0 and 0.8, and measurement rate ranged from one minute to one day. The node took

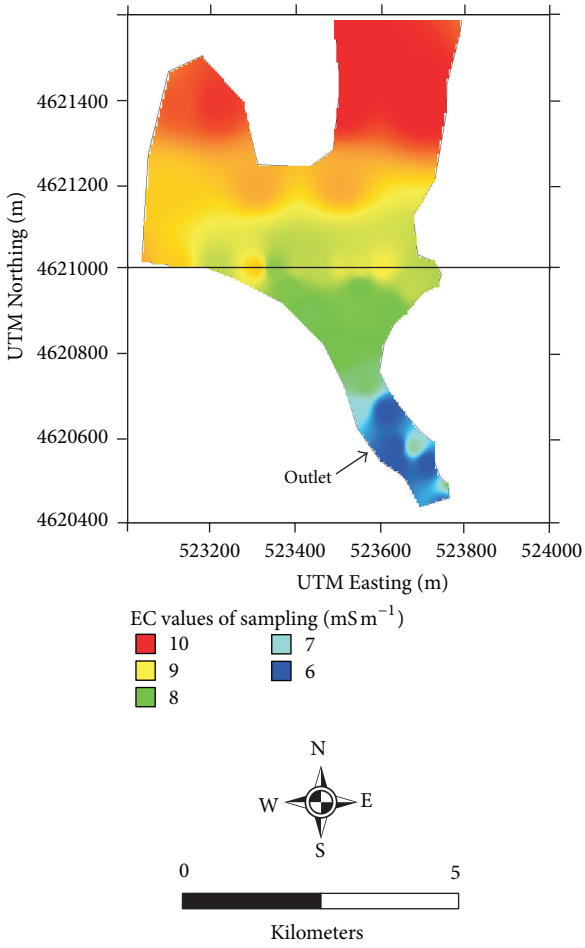


Figure 9: EC values of the sample dataset.

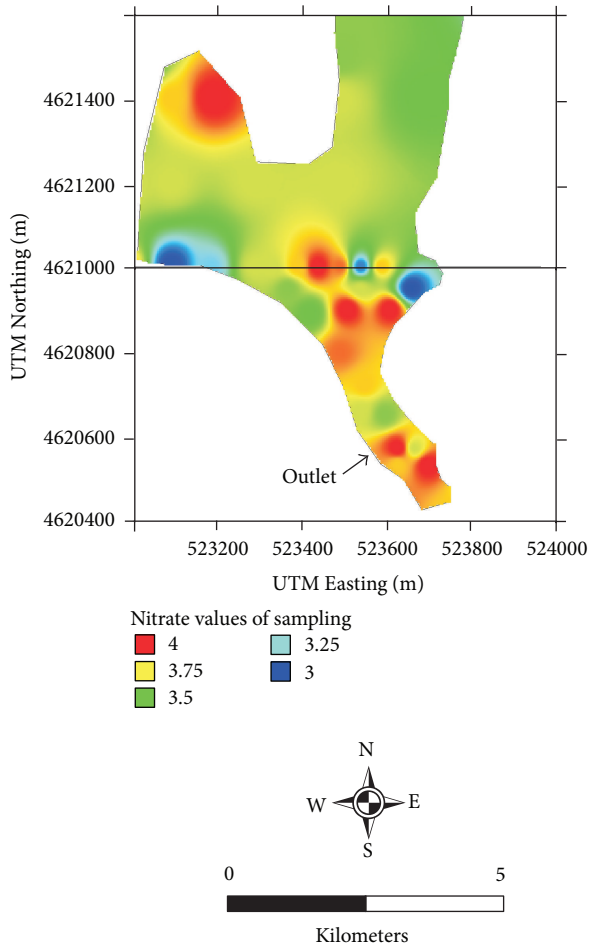


Figure 10: Nitrate values of the sample dataset.

40 samples out of the variable ranges and communicated with other nodes after processing the measurements 60% of the times. The resulting plot is shown in Figure 12. As in Figure 12, when transmission power is reduced, less power is used for the transmission mode of a radio. On the other hand, reducing transmission power reduces the transmission range of a radio and affects overall packet reception rate (PRR). Therefore, adaptive transmission power approaches may be utilized to balance this trade-off.

5.3. Future Research Directions. Two sets of field tests are going to be implemented at a drinking water reservoir—Kirklareli Dam, Kirklareli, Turkey. In the first set of field tests, a custom designed prototype boat will be programmed to use spiral paths. Because, although parallel paths are preferred in water quality mapping operations, spiral or other types of paths may be more appropriate to mitigate the temporal variations in the water quality parameters [13]. Table 3 lists the specifications of the boat [37]. We are planning to collect data regarding various water quality parameters including CTD, DO, pH, nitrate, and turbidity during the field tests using the autonomous boat. After processing and interpolating the collected data, pictures showing the levels of the water quality parameters of the Dam are going to be obtained.

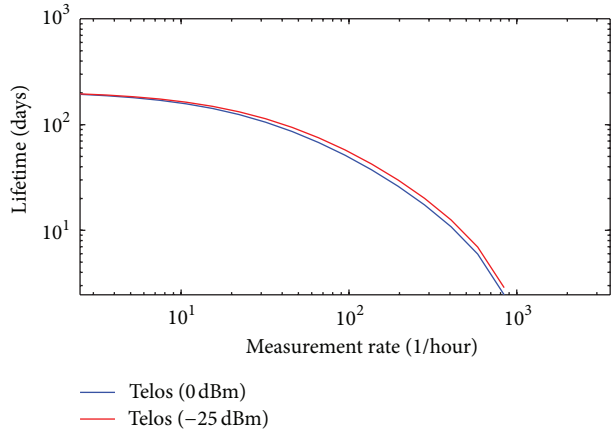


Figure 11: The relation between node lifetime and measurement rate.

In the second set of field tests, we are going to implement a WSN consisting of several nodes. These nodes are going to be buoys carrying sondes with multiple probes and have IEEE 802.15.4 wireless interfaces for communication. The buoys are going to be manually positioned at fixed locations at the Dam. They are going to analyze water quality at periodical intervals by using their probes and transmit their measurements to

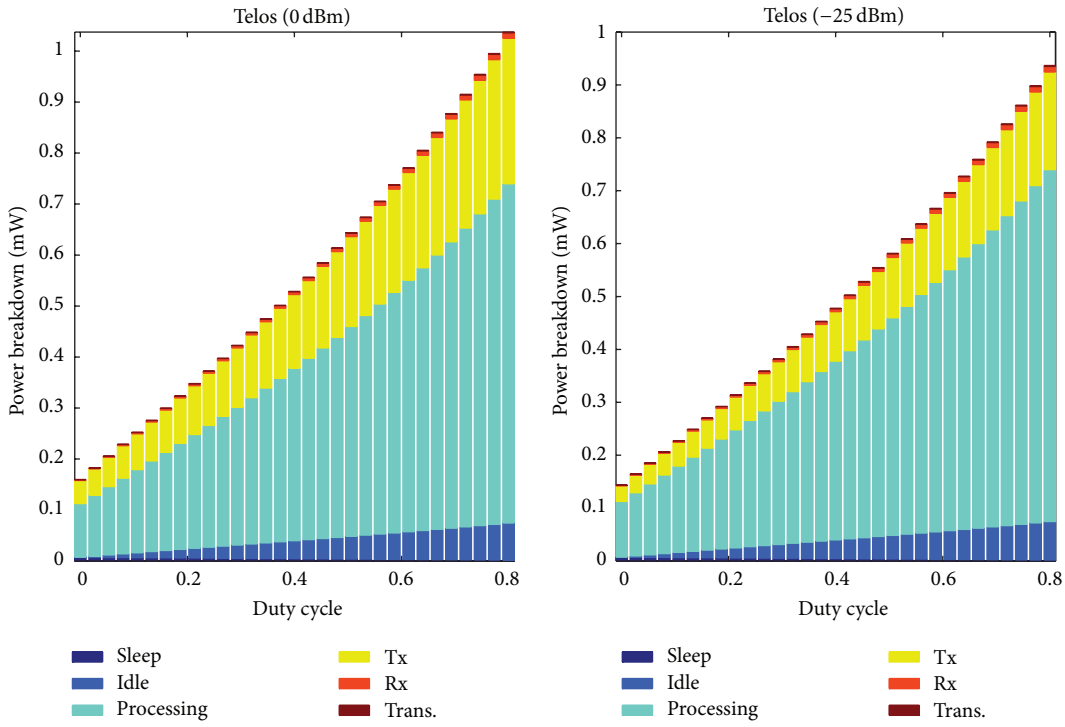


Figure 12: The relation between transmission power and battery consumption.

Table 3: Specifications of the custom designed boat.

Specification	Value
Hull material	UV Resistant ABS
Hull dimensions	180 × 90 cm
Weight	25 kg
Payload	30 lbs
Motor	Brushless DC outdrive
Typical speed	4 m/s
Maximum speed	5 m/s
Power	24 V NiMH Packs
Estimated Battery life—maximum speed	45 min.
Estimated Battery life—typical speed	45–140 min.
Depth sounder transducer	External mount
Communication	Two-channel radio modem to transmit boat position/heading/bathymetry and water quality data
Probes	CTD, DO, pH, turbidity, nitrate

the central PC over the WSN. In this way, continuous online monitoring of water quality is going to be realized.

A future design goal to be implemented on the prototype boat is to integrate solar and wind panels to improve battery life. As proven in several studies such as [38, 39], energy harvesting methods can considerably increase node battery life. In this way, longer lifetime and continuous use of the boat may be achieved by charging batteries when there is more power than needed. This design can be efficient if completed with additional sensors for different types of measurements including temperature, turbidity, and salinity.

Comparing the performance of the tight integration approach with that of the loose integration approach in terms of navigation accuracy, processing overhead, and memory usage is our final future work in this study. This future work aims to prove the applicability of GPS-INS integration approaches in practical implementations.

6. A Discussion on the Proposed Approaches

In this study, two different water quality monitoring systems for different scenarios are proposed. In the first system, an autonomous boat or a group of boats receive trajectories set by utility operators and then follow the trajectories to visit predefined sampling points. During the mission, the sondes on the boat(s) analyze water quality and log results. In the second system, a group of portable water quality probes mounted on buoys regularly analyzes water quality and sends the data to the utility control center through wireless or wired links.

Although the proposed approaches do not complement each other since they are different approaches, they can substitute each other depending on requirements. A comparison of the proposed approaches is listed in Table 4.

Table 4: WSN-based water quality monitoring versus water quality monitoring using an autonomous boat.

Method	WSN-based water quality monitoring	Water quality monitoring using an autonomous boat
Total time needed for deployment	Depends on deployment area (from a few hours to a few days)	N/A
Total time needed for analysis	Little (usually in a few minutes)	High (depends on deployment area)
Total cost of ownership including deployment	High (depends on the total number of Sondes and probe diversity)	Low (20 K–50 K)
Total number of measurements per analysis for a site	Medium to high (depends on the number of Sondes)	Medium to high (depends on the distance between paths followed)
The need for qualified operators	No	Yes
Real-time analysis	Yes	No
Using at multiple sites	Difficult	Easy
Probe diversity	Medium	High

7. Conclusions

This paper presents the design of two portable and low-cost solutions for continuous monitoring of water quality. Firstly, we have explained the design of an autonomous miniboat-aided autonomous water quality monitoring system for drinking water reservoirs. Then, we have investigated the use of portable sondes with IEEE 802.15.4-based wireless interfaces for continuous monitoring of water quality. These systems eliminate the need for periodical time-consuming water analyses and improve the quality of supplied water through continuous monitoring. In addition, they bring cost advantages to utility providers by eliminating periodical laboratory expenses.

The first system we have proposed utilizes an autonomous battery-operated miniboat which is guided by a GPS-aided INS-based navigation algorithm. Hence, the system can be used for water quality monitoring at multiple sites. The results of our first set of simulation studies prove that the proposed system is successful in terms of navigation and water sampling performance. The second system we have proposed utilizes a group of portable sondes with solar panels for energy harvesting and IEEE 802.15.4-based wireless interfaces mounted on buoys. The sondes form a WSN to communicate over and send their measurements at regular intervals to a central PC over the WSN. Similar to other WSN-based applications, the results of our second set of simulation studies show that the applicability of this system depends on several parameters such as transmission frequency, transmission power, packet size, and node-related parameters.

References

[1] T. G. Kazi, M. B. Arain, M. K. Jamali et al., "Assessment of water quality of polluted lake using multivariate statistical techniques: a case study," *Ecotoxicology and Environmental Safety*, vol. 72, pp. 301–309, 2008.

[2] C. Filik Iscen, O. Emiroglu, S. Ilhan, N. Arslan, V. Yilmaz, and S. Ahiska, "Application of multivariate statistical techniques in the assessment of surface water quality in Uluabat Lake, Turkey," *Environmental Monitoring and Assessment*, vol. 144, no. 1–3, pp. 269–276, 2008.

[3] "Directive 2000/60/EC. Water Framework Directive of the European Parliament and of the Council of 23 October 2000 establishing a framework for community action in the field of water policy," Official Journal L, 327, 2000, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32000L0060:EN:NOT.

[4] US EPA, "Groundwater and drinking water, drinking water standards," United States Environmental Protection Agency, 2003, http://www.epa.gov/safewater.

[5] SSKY, "Su KirliligiKontroluYonetmeligi. Water pollution control regulation," Turk Off Gaz 25687, 2004.

[6] J. Wang, L. Zhu, G. Daut et al., "Investigation of bathymetry and water quality of Lake Nam Co, the largest lake on the central Tibetan Plateau, China," *Limnology*, vol. 10, no. 2, pp. 149–158, 2009.

[7] R. J. Wilcock, R. M. Monaghan, J. M. Quinn et al., "Land-use impacts and water quality targets in the intensive dairying catchment of the Toenepi Stream, New Zealand," *New Zealand Journal of Marine and Freshwater Research*, vol. 40, no. 1, pp. 123–140, 2006.

[8] D. R. Lapen, E. Topp, M. Edwards et al., "Effect of liquid municipal biosolid application method on tile and ground water quality," *Journal of Environmental Quality*, vol. 37, no. 3, pp. 925–936, 2008.

[9] M. Rouen, G. George, J. Kelly, M. Lee, and E. Moreno-Ostoa, "High-resolution automatic water quality monitoring systems applied to catchment and reservoir monitoring," *Freshwater Forum*, vol. 23, pp. 20–37, 2004.

[10] H. B. Glasgow, J. M. Burkholder, R. E. Reed, A. J. Lewitus, and J. E. Kleinman, "Real-time remote monitoring of water quality: a review of current applications, and advancements in sensor, telemetry, and computing technologies," *Journal of Experimental Marine Biology and Ecology*, vol. 300, no. 1-2, pp. 409–448, 2004.

[11] H. T. Do, S. L. Lo, P. T. Chiueh, and L. A. P. Thi, "Design of sampling locations for mountainous river monitoring," *Environmental Modelling and Software*, vol. 27-28, pp. 62–70, 2012.

[12] P. E. Hagen, N. Størkersen, B. E. Marthinsen, G. Sten, and K. Vestgård, "Rapid environmental assessment with autonomous underwater vehicles: examples from HUGIN operations," *Journal of Marine Systems*, vol. 69, no. 1-2, pp. 137–145, 2008.

[13] Y. Kaizu, M. Iio, H. Yamada, and N. Noguchi, "Development of unmanned airboat for water-quality mapping," *Biosystems Engineering*, vol. 109, no. 4, pp. 338–347, 2011.

[14] X. Zhu, D. Li, D. He, J. Wang, D. Ma, and F. Li, "A remote wireless system for water quality online monitoring in intensive fish culture," *Computers and Electronics in Agriculture*, vol. 71, no. 1, pp. S3–S9, 2010.

[15] K. Singh, J. Worley, and L. M. Risse, "Sludge measurement using global positioning system (GPS) enabled sonar equipped airboat in a lagoon," *Applied Engineering in Agriculture*, vol. 24, no. 5, pp. 603–609, 2008.

[16] J. A. Lopez Riquelmea, F. Sotoa, J. Suardiaza, P. Sancheza, A. Iborraa, and J. A. Verab, "Wireless Sensor Networks for precision horticulture in Southern Spain," *Computers and Electronics in Agriculture*, vol. 68, no. 1, pp. 25–35, 2009.

[17] J. V. Capella, A. Bonastre, R. Ors, and M. Peris, "In line river monitoring of nitrate concentration by means of a Wireless Sensor Network with energy harvesting," *Sensors and Actuators B*, vol. 177, pp. 419–427, 2013.

[18] J. V. Capella, A. Bonastre, R. Ors, and M. Peris, "A Wireless Sensor Network approach for distributed in-line chemical analysis of water," *Talanta*, vol. 80, no. 5, pp. 1789–1798, 2010.

[19] R. Yue and T. Ying, "A novel water quality monitoring system based on solar power supply and wireless sensor network," *Procedia Environmental Sciences*, vol. 12, pp. 265–272, 2012.

[20] S. Zhuiykov, "Solid-state sensors monitoring parameters of water-quality for the next generation of wireless sensor networks," *Sensors and Actuators B*, vol. 161, no. 1, pp. 1–20, 2012.

[21] H. Erckens, G. A. Busser, C. Pradalier, and R. Siegwartand, "Navigation strategy and trajectory following controller for an autonomous sailing vessel," *IEEE Transaction on Robotics and Automation Magazine*, vol. 17, no. 1, pp. 45–54, 2010.

[22] N. Cruz and J. Alves, "Auto-Heading Controller for an Autonomous Sailboat," in *Proceedings of the IEEE Conference Oceans'10*, pp. 1–6, Sydney, Australia, May 2010.

[23] K. Xiao, J. Sliwka, and L. Jaulin, "A wind-independent control strategy for autonomous sailboats based on Voronoi diagram," in *Proceedings of the 14th International Conference on Climbing and Walking Robots and the Support Technologies for Mobile Machines*, 2012.

[24] R. Stelzer, T. Pröll, and R. I. John, "Fuzzy logic control system for autonomous sailboats," in *Proceedings of the IEEE International Conference on Fuzzy Systems*, July 2007.

[25] S. Godha and M. E. Cannon, "Integration of DGPS with a low cost MEMS: based Inertial Measurement Unit (IMU) for land vehicle navigation application," in *Proceedings of the 18th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS '05)*, pp. 333–345, September 2005.

[26] M. Petovello, *Real-time Integration of a Tactical-Grade IMU and GPS for High-Accuracy Positioning and Navigation [Ph.D. thesis]*, University of Calgary, Calgary, Canada, UCGE Report 20173, 2003.

[27] Y. Yang, *Tightly Coupled MEMS INS/GPS Integration with INS Aided Receiver Tracking Loops [Ph.D. thesis]*, University of Calgary, Calgary, Canada, UCGE Report 20270, 2008.

[28] E. D. Kaplan, *Understanding GPS Principles and Applications*, Artech House, Norwood, NJ, USA, 1996.

[29] J. A. Farrell and M. Barth, *The Global Positioning System and Inertial Navigation*, McGraw-Hill, New York, NY, USA, 1998.

[30] M. A. H. Franson, *Standard Methods for the Examination of Water and Wastewater*, American Public Health Association, Washington, DC, USA, 1989.

[31] ASTM D888-12, "Standard test methods for dissolved oxygen in water," ASTM International, West Conshohocken, Pa, USA, 2012, http://www.astm.org.

[32] US EPA, "YSI, Inc. 6600 EDS Multi-Parameter Water Quality Prove/Sonde, Joint Verification Statement," United States Environmental Protection Agency, 2003–2004, http://www.epa.gov/etv.

[33] http://atlas-scientific.com/sensors.html.

[34] http://www.libelium.com/products/waspmote.

[35] D. Jung, T. Teixeira, A. Barton-Sweeney, and A. Savvides, "Model-based design exploration of wireless sensor node lifetimes," in *Proceedings of EWSN*, pp. 277–292, Delft, The Netherlands, 2007 January.

[36] http://www.memsic.com/products/wireless-sensor-networks/wireless-modules.html.

[37] http://www.oceanscience.com/Products/Q-Boats/Q-Boat-1800D.aspx.

[38] G. Tuna, V. C. Gungor, and K. Gulez, "Wireless sensor networks for smart grid applications: a case study on link reliability and node lifetime evaluations in power distribution systems," *International Journal of Distributed Sensor Networks*, vol. 2013, Article ID 796248, 11 pages, 2013.

[39] G. Tuna, V. C. Gungor, and K. Gulez, "Energy harvesting techniques for industrial wireless sensor networks," in *Industrial Wireless Sensor Networks: Applications, Protocols, Standards, and Products*, G. P. Hancke and V. C. Gungor, Eds., pp. 119–136, CRC Press, 2013.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 602923, 11 pages
http://dx.doi.org/10.1155/2013/602923

Research Article

Deadline-Aware Scheduling Perspectives in Industrial Wireless Networks: A Comparison between IEEE 802.15.4 and Bluetooth

Mario Collotta, Giovanni Pau, and Gianfranco Scatà

Telematic Engineering Laboratory, Kore University of Enna, Cittadella Universitaria, 94100 Enna, Italy

Correspondence should be addressed to Mario Collotta; mario.collotta@unikore.it

Received 26 February 2013; Accepted 18 May 2013

Academic Editor: Tai-hoon Kim

In industrial contexts, most of process control applications use wired communication networks. The reliability of wired networks is indisputable and extensively demonstrated by several studies in the literature. However, it is important to consider several disadvantages provided by the use of wired technologies, like high deployment and maintenance costs and low network scalability. Although it is difficult to fully replace wired networks, wireless communication protocols have features which could undeniably affect in positive way the production mechanisms in factories. The wireless networks (WNs) are effectively used to detect and exchange information. The main communication protocols, currently available for WNs, however, do not support real-time periodic traffic flows which, as known, mainly characterize industrial networks. In this paper, we will analyze a real-time scheduling algorithm for both periodic and aperiodic traffic management, applied to networks based on IEEE 802.15.4 and Bluetooth, respectively. The main purpose of this research is to reduce, as much as possible, the packet loss on the channel, increasing at the same time the reliability of the wireless technology. Furthermore, the comparison between IEEE 802.15.4 and Bluetooth will allow to identify the more suitable communication protocol for industrial process control systems.

1. Introduction

The WNs are increasingly used in everyday life and in several heterogeneous application fields like home automation [1–5], road monitoring [6, 7], industry [8, 9], agriculture [10, 11], and health care [12, 13] to mention some. The WNs are characterized by several nodes connected to each other using architectures and protocols which depend on the environment where the application must be used. In different network topologies, nodes can work both as simple transmitters/receivers and as routers working in multihop mode. These environments are geographically limited but may be densely populated by nodes. This leads to a greater complexity in terms of communication channel reliability, which must be able to reduce, as much as possible, errors caused by collisions and interferences, and in terms of necessary costs for network implementation and management. As a consequence, a WN must satisfy several key features.

(i) *Reliability*. The communication mechanism adopted in a WN must guarantee maximum reliability during data transmission, avoiding collisions and therefore ensuring the integrity of data transmitted.

(ii) *Interoperability*. The devices used in a WN can be produced by different manufacturers. This must not represent an obstacle for devices integration inside the same WN. For this reason, the standard communication protocol adopted must be able to support the functioning of heterogeneous devices and must also allow the integration among different communication technologies, both wireless and wired.

(iii) *Scalability* [14, 15]. A WN must be able to dynamically manage the variation of the number of nodes of a network. Some devices, in fact, can be of mobile. The network must be able to configure itself and independently manage the network topology.

(iv) *Low Cost*. In order to make convenient the wireless technology adoption, it is necessary to make competitive devices prices. During last years, the reduction of devices prices has made "low cost" this type of technology.

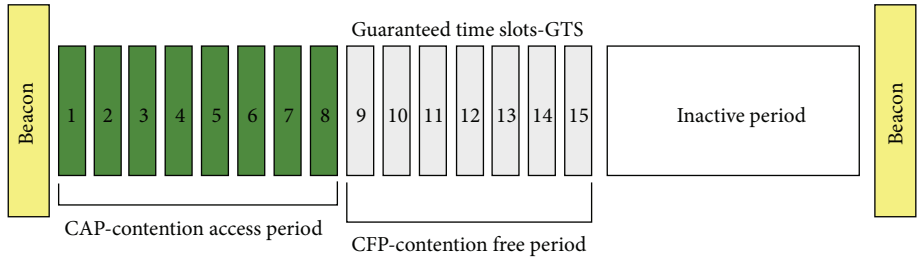


Figure 1: IEEE 802.15.4 MAC superframe structure.

(v) *Low Data Rate*. In a WN, devices do not need very high data rates. This is due to the intrinsic characteristics of data to manage.

(vi) *Low Power* [16]. The WNs are often deployed in environmental contexts where it is really difficult to access. So, it is very important to reduce power consumption. The WN must apply energy harvesting policies in order to increase devices autonomy and the lifecycle of the whole network.

Most important standard protocols for WNs communication are IEEE 802.15.4 [17], IEEE 802.15.1 Bluetooth [18], 6LoWPan [19], and WilressHART [20]. Among these protocols, IEEE 802.15.4/ZigBee and IEEE 802.15.1 Bluetooth have been extensively explored. The main problem of these protocols is represented by the fact that they do not implement deadline-oriented scheduling algorithms. Thus this involves inefficiency in soft real-time contexts like, for example, industrial process control systems. The IEEE 802.15.4 standard protocol can be used in soft real-time contexts through the Guaranteed Time Slots (GTS) mechanism provided by the "beacon enabled" mode. The standard provides the use of a superframe structure defined by PAN Coordinator. The superframe is delimited by special signalling packets, called beacons, and it is divided into 16 time slots (Figure 1).

All superframes have the same length and are characterized by an active period and an inactive period during which the PAN Coordinator and the devices can switch off the radio and enter in an energy saving state; however, the PAN Coordinator stays switched on for all active period duration. The active period is further divided into a Contention Access Period (CAP) followed by a Contention Free Period (CFP) without contention consisting of a variable number of continuous GTSs. The Bluetooth technology uses a master/slave protocol through which it is possible to connect up to seven active slaves for each wireless cell, better known as piconet, and adopts a combination of Frequency Hopping and Time Division Multiplexing (FH-TDM). Bluetooth has several characteristics which make it suitable for many application areas. The access to the transmission medium is managed by the master and it is more suitable for deterministic data transmissions because it is "*time slotted*." In other words, each slot has a fixed length (625 μs) as shown in Figure 2.

Moreover, this medium access mechanism is really efficient since it requires only 1250 μs for data exchange. Reduced size of each time slot produces some advantages for industrial applications since it allows to have very short time-cycles and

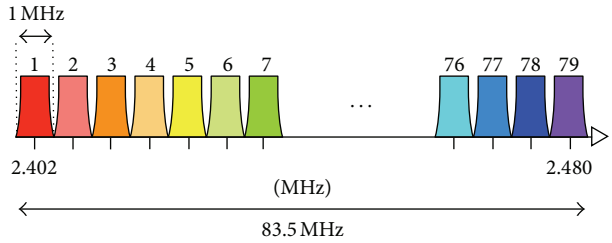


Figure 2: Bluetooth radio links.

higher accuracy (fundamental requirements in periodic real-time applications). Furthermore, The Frequency Hopping (FH), the Time Division Multiplexing (TDM), the Cyclic Redundancy Check (CRC), the Header-Error-Check (HEC), and the Forward Error Correction (FEC) allow the realization of a reliable and robust channel, in other words suitable for time-critical data transmission [21]. It is important to underline that Bluetooth does not natively support real-time communications. As demonstrated by [22], a deadline-aware scheduling should be introduced in order to ensure the satisfaction of real-time constraints. Moreover, the energy consumption of Bluetooth is not negligible although this does not represent matter of concern in industrial environments where automation applications usually require working cycles which prevent the adoption of the energy saving policies. Regarding the number of nodes in a piconet, seven slaves are enough for several applications at field level, in which each BT node can collect data from multiple devices, thus increasing the amount of exchanged information. In addition, several piconets can coexist and operate in parallel, thanks to the combination of the large number of available channels and to the frequency hopping, which allows the simultaneous presence of more transmissions on multiple channels [23]. However, in applications where there is a high number of devices and then a considerable amount of data exchange, more overlapping piconets are necessary.

This work shows an innovative deadline-aware scheduling approach for industrial process control scenarios. This approach is based on the combined use of EDF (for periodic real-time traffic flows scheduling) and CBS (for aperiodic traffic flows scheduling). Specifically, this paper shows an implementation of this approach both on IEEE 802.15.4 networks and on Bluetooth networks in order to determine the best technology meeting the requirements which characterize

time-critical industrial environments. The paper is organized as follows. Section 2 describes main related works in order to determine the current state of art. Section 3 describes the network architecture and the proposed approach showing its implementation both in IEEE 802.15.4 and in Bluetooth networks. Section 4 presents a test-bed scenario in order to show network performance obtained by the proposed approach in both network technologies. Finally, Section 5 summarizes the paper reporting conclusions.

2. Related Works

2.1. Techniques to Support Real-Time Transmissions over IEEE 802.15.4. In the beacon enabled mode, the IEEE 802.15.4 standard protocol allows nodes, characterized by real-time constraints, to allocate GTSs during the Contention Free Period. The protocol supports the assignment of GTSs; in other words, a node has a number of time slots for exclusive use on each superframe. The limitation of this explicit GTSs allocation is that the resources can run out quickly because in every superframe up to seven GTSs can be allocated. In this way, other nodes cannot use the guaranteed service. Furthermore, the CFP can be underused producing, as a consequence, bandwidth waste. In order to overcome these limitations, some researcher proposed i-GAME [24], through which it is possible to provide implicit GTSs allocation for IEEE 802.15.4 networks. The allocation is based on implicit GTS allocation requests, taking into account the traffic characteristics and delay requirements of the flows. The approach allows the use of a GTS by the nodes in a multiple way, ensuring that they always meet delay and bandwidth requirements. In [25] an algorithm is shown for admission control that allows to decide whether to accept a new assignment request GTS. This algorithm is not only based on the remaining time interval, but also on traffic flows specifications, on admissible delay and available bandwidth resources. This approach improves the bandwidth allocation mechanism of the IEEE 802.15.4 standard protocol. The GTS allocation concept is similar to the Time Division Multiplex Access (TDMA). Some reserved bandwidth is periodically guaranteed for a specific data flow. The bandwidth is determined taking into account the time window duration and its periodicity. However, the GTS mechanism is more flexible than the classical TDMA since the GTS duration can be dynamically adjusted by appropriately setting the superframe parameters of IEEE 802.15.4, while a TDMA slot has a time duration generally fixed for a given network configuration. Furthermore, the TDMA and the IEEE 802.15.4 MAC layer differ in several aspects. In other words, the IEEE 802.15.4 presents more advantages than TDMA in WNs applications. The scalability is the most significant limitation of the TDMA-based approach, since the number of nodes in a TDMA cluster must be kept as low as possible. This prevents its use in WNs, while the IEEE 802.15.4 could manage up to 254 nodes in a cluster. Another problem of TDMA is the lack of support for dynamic network topology changes (e.g., node failure, a new node enters inside the network, mobility), since in the TDMA the network configuration must be readapted each time the topology changes. Moreover, communications in TDMA-based networks are quite dependent on the cluster manager. If this node fails, the underlying sensor nodes will be disconnected from the entire network. This is not the case of IEEE 802.15.4, because the protocol can work with or without a central manager, and it is designed to be easily adapted to different network topologies. The performance of the IEEE 802.15.4 protocol has been subject of several research studies, focused mainly on CSMA/CA performance (Carrier Sense Multiple Access with Collision Avoidance) [26, 27] also using simulations [28]. In recent years, researchers focused on real-time systems and Quality of Service (QoS) problems in WNs. An overview of these challenges is presented in [29]. An example of TDMA adaptation is extensively discussed in [30]. In this work, an optimization method for dynamic sharing of time slots has been proposed. However, this protocol is based on the next node detection through a planned exchange of information. As a consequence, this method produces a considerable overhead. Another variant of adaptive TDMA is shown in [31] where the management of virtual time slots for control messages is based on CSMA. Some existing hybrid protocols make a distinction between the contention access period (CAP) and contention free period (CFP) to avoid collisions. However, reservation request for a GTS can be made through a message in the CAP using the CMSA/CA. Consequently, the reservation request depends on the use of the current channel and can be delayed for a time interval unknown.

2.2. Techniques to Support Real-Time Transmission over Bluetooth. The BT Master/Slave (M/S) approach is often inefficient in process control applications that require a regular data exchange. In addition, the fixed overhead, introduced into the slot used by the Master for polling, causes a waste of bandwidth that can degrade system performance. For this reason in the literature several works proposing the use of the Slave/Slave (S/S) communication have been proposed. In [32] the authors propose an approach in which M/S switching and the piconet partitioning are triggered when a Slave needs to communicate frequently with another Slave. An improvement of this approach concerning the switching bottleneck is presented in [33]. However, these approaches enable only the communication between pairs of Slaves, thus restricting the possible application scenarios. An innovative approach that allows Slaves to communicate with any other member of a group of Slaves is presented in [34]. This feature makes the approach highly suitable for data exchange in Distributed Process Control Systems (DPCSs). The Master configures the Slaves in order to make them belong to a logical ring, and in addition, it specifies the order with which they can transmit. Each Slave is able to communicate with the other Slaves without passing through the Master. In this way the overhead is significantly reduced. Moreover, the compatibility with the standard BT M/S mode is guaranteed because the nodes can operate in both modes. Therefore significant changes to the hardware level are not required. In the initial phase, the Master is able to obtain all scheduling necessary information. This is possible because the DPCS applications generally have known dynamic. Then the Master also defines which nodes must operate in S/S mode and when they can do it. Moreover,

the Master decides whether or not to activate parallel M/S and S/S communication. According to the Master's scheduling order, the Slaves, belonging to the ring logical, can transmit when they have a virtual token. In the BT standard, the use of fixed length slots allows to keep synchronization between various nodes of the network and it prevents the overlapping between two consecutive transmissions. Through the local scheduling scheme, each node can transmit to each active Slave in the logical ring. In a piconet organized in this way, the S/S transmission sequence can be repeated several times. Since the stations are not synchronized with the Master during S/S communications, the number of repetitions should be at least less than the maximum time over which the synchronization with the Master is loosed. This maximum value has been calculated analytically in [34] and it is approximately equal to 800 slots. These characteristics show that it is possible to generate long S/S transmission sequences without losing synchronization. When the time allowed for the logical ring expires, the last Slave synchronizes with the hopping frequency of the master. From the next slot, the original piconet is restored and the Master can resume the transmission with all Slaves in M/S mode. Periodic and aperiodic traffic scheduling in low voltage networks is analyzed in [35]. In one of the proposed approaches the periodic traffic transmission of each Slave is managed by the Master through the Earliest Deadline First algorithm [36] while the aperiodic traffic is managed through a Total Bandwidth Server (TBS) [36]. In any case, this approach is affected with the limitations mentioned earlier such as scan cycle duration and waste of bandwidth, since it is based on standard BT M/S mode. Regarding BT, in our approach we propose a novel real-time deadline-aware scheduling in S/S mode communication for an integrated management of periodic and the aperiodic traffic flows. The proposed approach is based on EDF and CBS [36], a dynamic priority server.

3. WNs-Based Solutions for Real-Time Networks

3.1. Deadline-Aware Scheduling Techniques in Wireless Networks. In this paper, we propose a deadline-aware scheduling for both periodic and aperiodic/sporadic traffic flows. Periodic flows (each one with a period T_i and a Computational Time C_i) are scheduled using the Earliest Deadline First (EDF) algorithm [36], while aperiodic and sporadic flows are managed through the Constant Bandwidth Server (CBS) mechanism [36]. EDF handles scheduling requests assigning a priority that is inversely proportional to the absolute deadlines. In our approach the scheduling requests refer to periodic variables transmission requests of the controlled industrial process. Given a periodic transmission request that arrives at time r_i, with relative deadline D_i (which is the time interval measured from the arrival time of the request to the time within which it must be transmitted) equal to the period T_i, the absolute deadline is shown in

$$d_i = r_i + D_i. \tag{1}$$

It is important to underline that, as demonstrated by Liu and Layland [37], in case of tasks' scheduling, a set of periodic tasks is schedulable with EDF if and only if the total utilization factor (U) is ≤1 as described by

$$U = \left(\sum_{i=1}^{n} \frac{C_i}{T_i} \right) \leq 1, \tag{2}$$

where C_i is the transmission duration and T_i is the period of the ith transmission.

This equation can also be used in this context where, obviously, we will have a nonpreemptive scheduling.

In the literature the CBS is defined through two parameters: the period T_s and the capacity Q_s defined as the maximum budget usable by a request at any period T_s. The equation

$$U_s = \frac{Q_s}{T_s} \tag{3}$$

represents the server bandwidth. The server also manages two internal variables in order to define its status; c_s represents the current budget at period t, while d_s is the current deadline assigned by the server to a request. These values are initialized to zero. To each aperiodic request, a budget Q_s, equal to the server capacity, and a starting relative deadline, equal to the server period T_s, are assigned. When an active request is completed, the network coordinator controls if the residual server budget is enough to manage other requests. If the budget is enough, the network coordinator schedules the first element in the queue with current deadline and capacity. When $c_s = 0$, the following conditions are considered:

$$c_s = Q_s, \tag{4}$$

$$d_s = d_s + T_s. \tag{5}$$

In other words, the capacity is supplied with the maximum value Q_s and the current deadline is put behind of a period T_s. In addition, the network coordinator controls if it is possible to recycle the server's budget and the current deadline when, in case of server idle, a request arrives at time t. The network coordinator verifies the following condition:

$$c_s < (t_s - D_s) * \left(\frac{Q_s}{T_s} \right). \tag{6}$$

If condition (6) is verified, the request can be scheduled with the current server's values. Otherwise, the c_s value must be reset to the maximum value Q_s, according to (4), and the deadline recalculated according to

$$d_s = t + T_s. \tag{7}$$

In order to guarantee messages schedulability, in accordance with deadlines, it is possible to use the known Jeffay's theorem [36], since we are in a non-preemptive scheduling context. A set of periodic requests (messages) is schedulable, using a non-preemptive algorithm, if two conditions are met. The first (8) relates to system utilization (in terms of bandwidth, as we are dealing with the transmission of packets), whereas the second (9) refers to the system demand.

Theorem 1. *A system can schedule a set of periodic requests, using non-preemptive EDF algorithm, if Jeffay's conditions (*(8) *and* (9)*) are met as follows:*

$$U_{\text{tot}} = U_p + U_s = \left(\sum_{i=1}^{n} \frac{C_i}{T_i} + U_s \right) \leq 1, \tag{8}$$

$$T_1 < L < T_n : L \geq C_i + \sum_{j=1}^{i-1} \left\lfloor \frac{L-1}{T_j} \right\rfloor C_j, \quad 1 < i \leq n; \ \forall L. \tag{9}$$

The periodic traffic flows are represented by a set of periodic variables $\tau_p = \{p_1, p_2, \ldots, p_n\}$*, where* $p_i = (C_i, T_i)$*, sorted in nondecreasing order by period (i.e., for any pair of variables* p_i *and* p_j*, if* $i > j$*, then* $T_i \geq T_j$*), and* C_i *is the transmission time for a periodic traffic flow generated by i*th *wireless node.*

Equation (8) relates to system utilization (in terms of bandwidth, as we are dealing with the transmission of packets), whereas (9) refers to the system demand. Equation (8) defines that total bandwidth utilization must not exceed 1; U_p is the utilization factor for periodic traffic while U_s is the utilization factor for sporadic and aperiodic traffic flows (i.e., server utilization). The inequality in (8) refers to a least upper bound on bandwidth demand that can be achieved in an interval of length L. This interval starts when the periodic variable is invoked and ends before the relative deadline. Then, a set of variables is schedulable if the demand in the interval L is less than or equal to the length of the interval.

3.2. EDF + CBS in IEEE 802.15.4. In IEEE 802.15.4 networks, the CBS can be implemented, for example, by choosing to use the solution proposed by Collotta et al. [23]. As mentioned in Section 2, the authors propose a solution, called i-GAME, where multiple nodes can share a GTS. In this case, the nodes which transmit aperiodic/sporadic traffic flows will use the GTSs sharing technique described in [24], which involves the use of a GTS by multiple nodes guaranteeing, to each one, the satisfaction of real-time constraints. So, in our approach, a message has the highest priority if its deadline is the closest among all those relating to messages ready to be transmitted. In this way, the message transmitted is always the one with more imminent deadline. In order to guarantee messages schedulability, following the Jeffay's conditions and the EDF algorithm, the total utilization factor for periodic traffic flows must be

$$U_{\text{tot}} = U_p + U_s = \frac{C_{\text{BF}} + C_{\text{CAP}} + C_{\text{CFP}} + C_{\text{IP}}}{\text{superframe}_{\text{duration}}} + U_s = \left(\sum_{i=1}^{n} \frac{C_i}{T_i} + U_s \right) \leq 1, \tag{10}$$

where, considering the superframe structure shown in Figure 1, C_{BF}, C_{CAP}, C_{CF}, and C_{IP} are, respectively, the beacon frame, CAP, CFP, and Inactive Period durations, respectively. According to (8) and (10), the utilization factor for aperiodic traffic flow can be obtained as follows:

$$U_s = 1 - U_p \tag{11}$$

$$U_s = 1 - \frac{C_{\text{BF}} + C_{\text{CAP}} + C_{\text{CFP}} + C_{\text{IP}}}{\text{superframe}_{\text{duration}}}. \tag{12}$$

3.3. EDF + CBS in Bluetooth. The approach consists of two scheduling levels. The first level is the local scheduling that it is carried out in the local queues of each BT device. In order to manage real-time traffic, we assume that in the local scheduling each device is managed by EDF. The second level, intrapiconet scheduling, refers to deadline-aware scheduling policies adopted inside the piconet, by the Master node. As in a typical industrial communication scenario, we assume that traffic exchanges are known a priori. On the contrary, regarding to the aperiodic traffic, we consider a signaling scheme where the Slaves send the status of their queue to the Master through a specific packet field. For simplicity we assume that each Slave transmits an aperiodic variable and a periodic variable, with fixed period and deadline equal to the period. In addition, each Slave manages different queues for aperiodic and periodic variables. However, our assumption is not mandatory since it could be possible to implement several distinct traffic classes if several different periodic and aperiodic variables are produced by each Slave. In addition, it is possible to use an S/S communication approach as described in [22]. This approach can be integrated to the functions already present in the BT standard. In fact, the development of new functions does not require hardware changes but firmware integrations. The approach shown in [22] called M/S (EDF+CBS)+S/S (EDF) provides an integrated support for deadline-aware scheduling of periodic and aperiodic traffic in parallel M/S and S/S communications. Specifically, S/S communications relate only to the periodic traffic while M/S transmissions handle hybrid stations which generate both periodic and aperiodic traffic. In our approach, we define a scheduling in a time period called Scan-cycle that is equal to the least common multiple of all traffic flows periods generated by the Slaves. The utilization factor for periodic traffic flow in M/S mode is obtained as follows:

$$U_{p_{\text{M/S}}} = \sum_{i=1}^{n} \frac{\text{SLOT}_{\text{DURATION}} + M_p}{T_i}, \tag{13}$$

where M_p is the Master polling slot (1 slot $= 625\,\mu$s) and n represents the number of Slaves ($n = 7$ for simplicity). Furthermore, considering that the M/S operating mode manages mixed traffic (periodic and aperiodic), U_s, previously defined as the utilization factor of the server that manages aperiodic flows, can be written according to

$$U_s = \frac{\text{APERIODIC}_{\text{TRAFFIC}_{\text{SLOT}_{\text{DURATION}}}} + M_p}{T_s}. \tag{14}$$

Considering (11) and assuming that each aperiodic variable can use a maximum of 5 transmission slots, (14) can be rewritten as follows:

$$1 - U_p = \frac{5+1}{T_s} = \frac{6}{T_s}, \tag{15}$$

from which it is possible to obtain the server period

$$T_s = \frac{6}{1 - U_p}. \tag{16}$$

Instead, as previously mentioned, the S/S mode manages only the periodic traffic flows. In this case the utilization factor for periodic traffic flow in S/S mode is obtained as follows:

$$U_{p_{S/S}} = \sum_{i=1}^{n} \frac{\text{SLOT}_{\text{DURATION}}}{T_i}. \tag{17}$$

As shown in (17), the polling slot of the Master is not present because it is not necessary. However, the Master may intervene in the communication in order to synchronize the Slaves after n Scan-cycle. Knowing the dynamics of periodic variables produced by various Slave, the Master through EDF carries out the polling in subpiconets through M/S mode and, afterwards, it schedules the S/S transmission in the logical ring. When the Slaves must transmit aperiodic traffic, they signal their aperiodic transmission request during the periodic packets transmission. After receiving this request, the Master manages it through CBS policies. In fact a budget, equal to the server capacity Q_s, and a relative deadline, equal to the server period T_s, are assigned to each aperiodic request. In order to manage all requests, the Master has to check the remaining budget of the server and, in the case in which it is equal to zero, the conditions shown in (4) and (5) must be carried out. In addition, for a better management of the requests coming to the server, even when it is idle, the Master must take into account the conditions shown in (6) and (7). The parallel use of the M/S and S/S communications is able to support higher workloads and it ensures more available bandwidth compared to the case where all nodes operate in M/S mode. In addition, in the S/S mode, the Master polling is not necessary and then the Scan-cycle for a control loop can be reduced significantly. A shorter Scan-cycle is able to improve the real-time traffic schedulability. Figure 3 shows the standard BT M/S communication while Figure 4 shows the shortened Scan-cycle obtained through the proposed parallelism between the M/S and S/S modes. The two subpiconets use different hopping frequencies because, as demonstrated in [23], it is possible to obtain a parallelism among piconets without cochannel interference problems. A further advantage of the S/S mode is that it is able to greatly reduce transmission/reception overhead introduced by correctness check, buffering time, packet assembly, and so forth. Instead, in the M/S mode, there is no overhead because the S/S communication is direct.

4. Performance Evaluation

In order to validate benefits introduced by the use of EDF + CBS deadline-aware scheduling in IEEE 802.15.4 and Bluetooth, several simulations have been carried out. Regarding

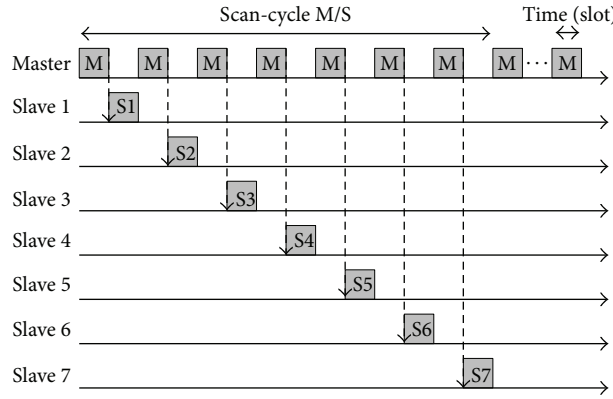


Figure 3: M/S communication.

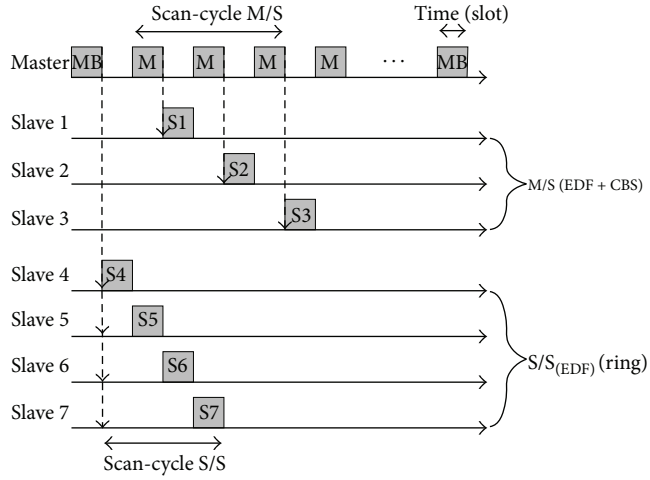


Figure 4: Parallel M/S and S/S communication.

IEEE 802.15.4, the simulations have been conducted using OMNet++ [38] considering a star topology consisting of a First Pan Coordinator (FPC) and seven end devices (RFD), as shown in Figure 5. The same topology has been used for measurements concerning a Bluetooth network (a master and seven slaves). In this case, simulations have been carried out through the ucbt extention of NS-2 simulator [39]. In both cases, the Throughput/Workload percentage, the deadline miss ratio, and delays related to both periodic and aperiodic flows have been measured considering packet size = 18 KB and data rate = 180 kbps for each station.

4.1. Performance Evaluation in an IEEE 802.15.4 Network. Figure 6 shows IEEE 802.15.4 network performance in terms of TH/WL ratio percentage for both periodic and aperiodic packets. As it is possible to see, the proposed approach produces a great improvement. The TH/WL ratio of periodic flows reaches 82.5% while the standard reaches 61.8%. Aperiodic flows reach 11.5% using EDF + CBS against the 28.2% obtained in case of standard protocol utilization. This happens because, in soft real-time environments, most of the traffic is periodic. So it is important to manage periodic flows as best as possible.

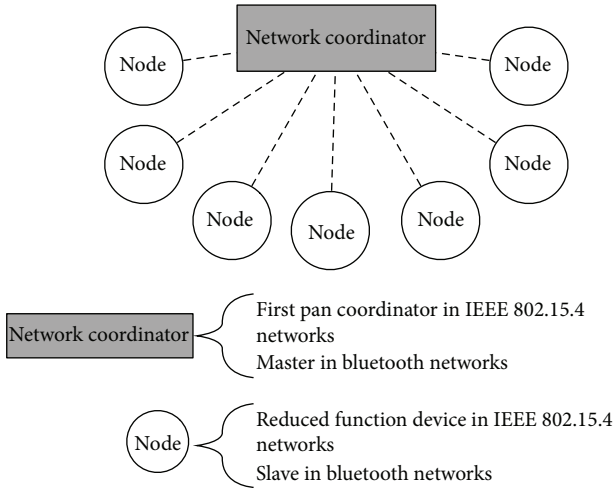


Figure 5: Network topology simulation scenario.

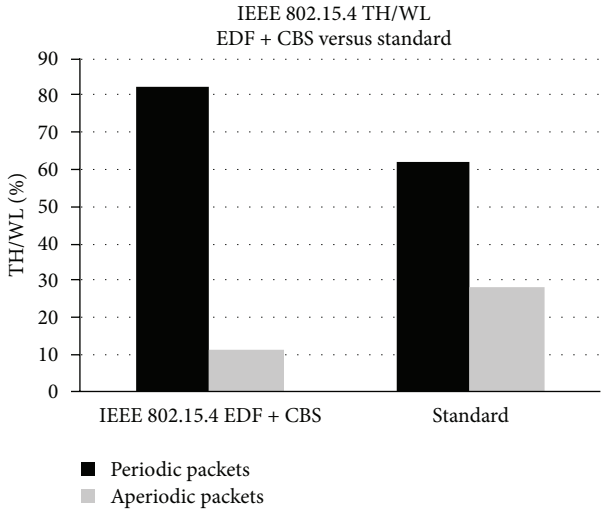


Figure 6: TH/WL—IEEE 802.15.4 Standard versus IEEE 802.15.4 EDF + CBS.

Figure 7 shows IEEE 802.15.4 network performance in terms of Deadline Miss ratio percentage for both periodic and aperiodic packets. In this case the proposed approach produces a reduction in terms of deadline miss. During simulations, 4.1% has been measured for periodic flows against 6.9% measured in case of standard protocol. Even in case of aperiodic traffic flows a reduction of deadline miss ratio has been measured (1.9% versus 3.1%).

Figures 8 and 9 show IEEE 802.15.4 network performance in terms of average delay of periodic and aperiodic packets, respectively. In x-axis, the packets sent percentage is represented while the y-axis measures the average delay in ms. Even these measures demonstrate benefits of the EDF + CBS approach. The delay of periodic traffic flows is on average 297.58 ms lower than the standard while the delay of aperiodic traffic flows is on average 331.09 ms lower than the standard.

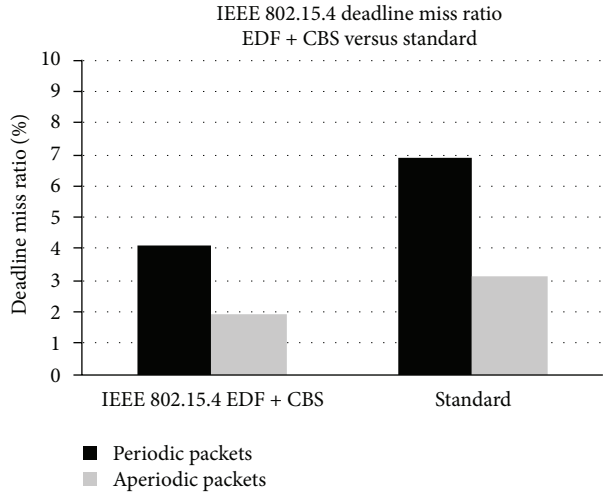


Figure 7: Deadline Miss ratio—IEEE 802.15.4 Standard versus IEEE 802.15.4 EDF + CBS.

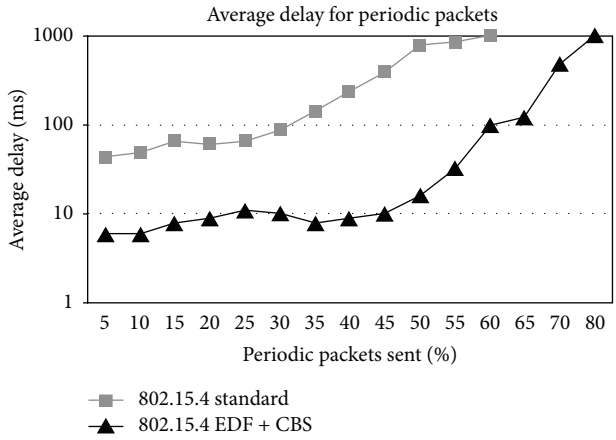


Figure 8: Average delay for periodic packets—IEEE 802.15.4 Standard versus IEEE 802.15.4 EDF + CBS.

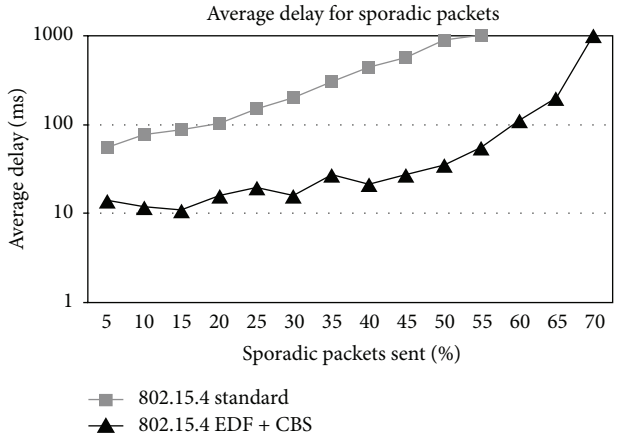


Figure 9: Average delay for aperiodic/sporadic packets—IEEE 802.15.4 Standard versus IEEE 802.15.4 EDF + CBS.

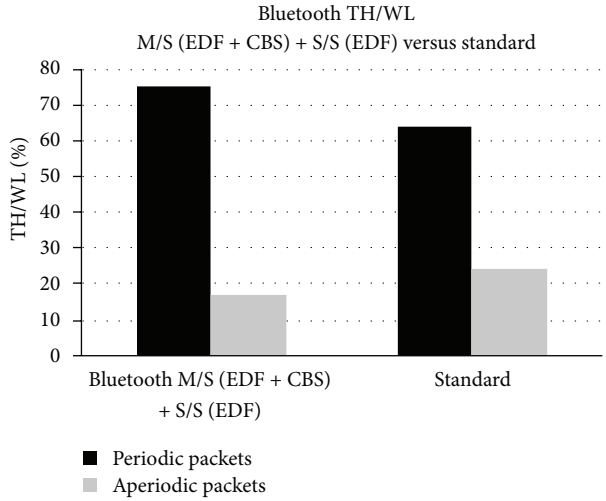


Figure 10: TH/WL—Bluetooth Standard versus Bluetooth M/S (EDF + CBS) + S/S (EDF).

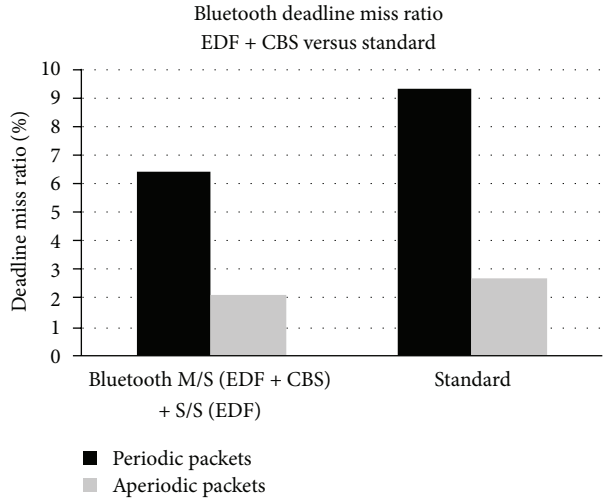


Figure 11: Deadline Miss ratio—Bluetooth Standard versus Bluetooth M/S (EDF + CBS) + S/S (EDF).

4.2. Performance Evaluation in a Bluetooth Network. Bluetooth network performance, in terms of TH/WL ratio percentage for both periodic and aperiodic packets, is shown in Figure 10. Even in this case, there are great improvements through the proposed approach. Measured TH/WL of periodic flows is 74.8% using M/S (EDF + CBS) + S/S (EDF) approach while the value obtained through standard protocol is 63.8%. On the contrary, TH/WL ratio value of aperiodic flows is 11.5% against 28.2% obtained in case of standard protocol utilization. As previously explained, in soft real-time environments most of the traffic is periodic.

Figure 11 shows Bluetooth network performance in terms of Deadline Miss ratio percentage for both periodic and aperiodic packets. Even in this case there is a deadline miss reduction using the proposed approach. A Deadline Miss ratio of 6.4% has been measured for periodic flows using the proposed approach while the standard protocol produces a value of 9.3%. Regarding aperiodic traffic flows, Figure 11 shows a Deadline Miss ratio reduction using our approach (2.1%) compared to the standard protocol (2.7%).

Figures 12 and 13 show Bluetooth network performance in terms of average delay of periodic and aperiodic packets, respectively. Even in this case, results show benefits of the proposed approach. In fact, the delay measured is on average 292.16 ms lower than the standard protocol in terms of periodic traffic flows. The delay of aperiodic traffic flows is, instead, on average 308.80 ms lower than the standard.

4.3. Performance Comparison between IEEE 802.15.4 and Bluetooth. In order to provide a comparison between IEEE 802.15.4 and Bluetooth solutions, performance obtained has been compared. Previously, we demonstrated that a deadline-aware solution improves performance of the standard protocol. Now, we want to summarize obtained results in order to determine the best solution for soft real-time applications. Figure 14 shows a comparison between results obtained in IEEE 802.15.4 and Bluetooth, respectively. In an IEEE

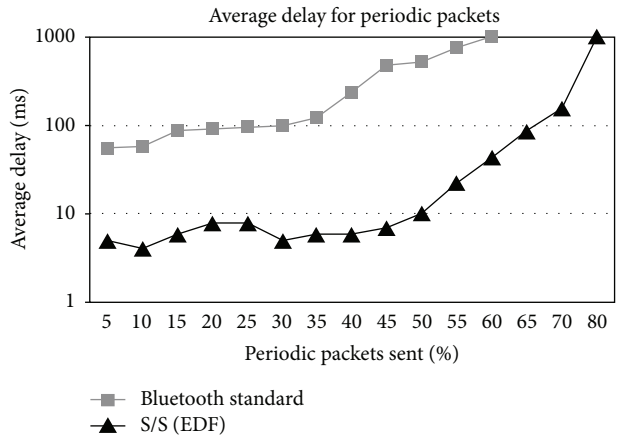


Figure 12: Average delay for periodic packets—Bluetooth Standard versus Bluetooth S/S (EDF).

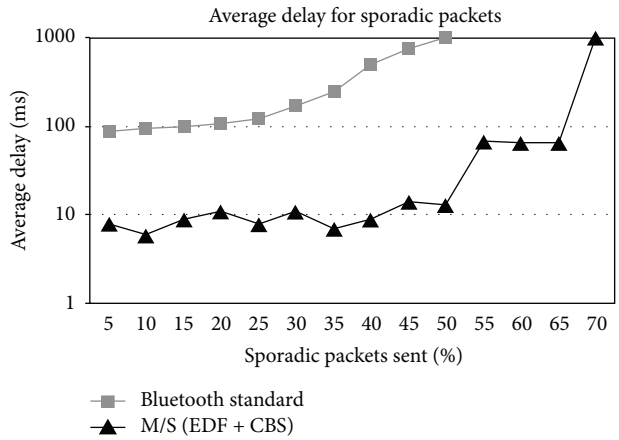


Figure 13: Average delay for aperiodic/sporadic packets—Bluetooth Standard versus Bluetooth M/S (EDF + CBS).

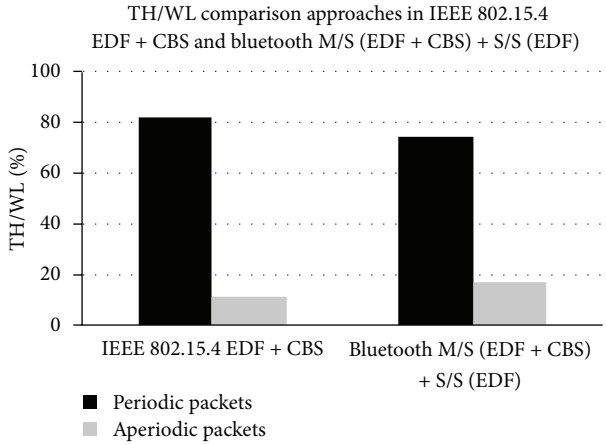


Figure 14: TH/WL—IEEE 802.15.4 EDF + CBS versus Bluetooth M/S (EDF + CBS) + S/S (EDF).

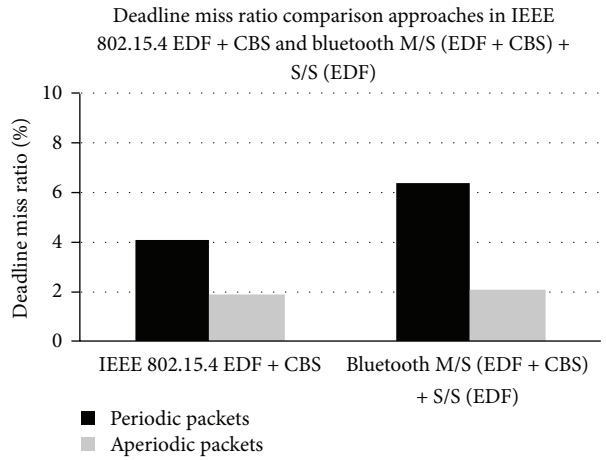


Figure 15: Deadline Miss ratio—IEEE 802.15.4 EDF + CBS versus Bluetooth M/S (EDF + CBS) + S/S (EDF).

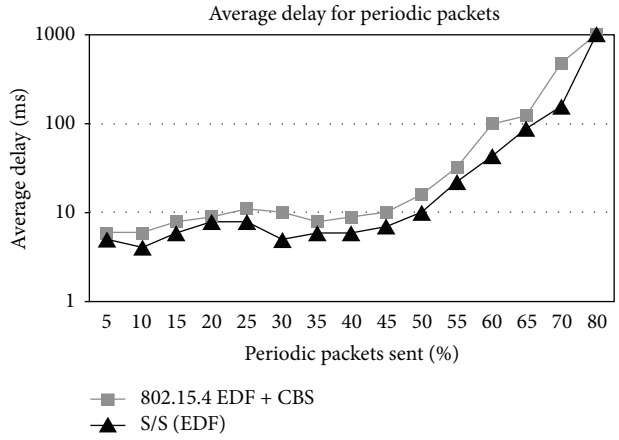


Figure 16: Average delay for periodic packets—IEEE 802.15.4 EDF + CBS versus Bluetooth S/S (EDF).

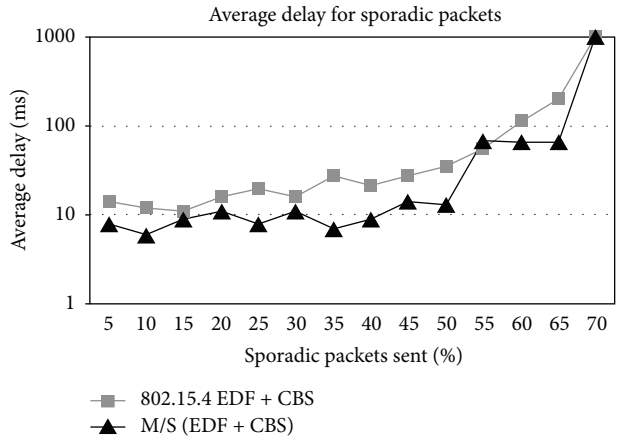


Figure 17: Average delay for aperiodic/sporadic packets—IEEE 802.15.4 EDF + CBS versus Bluetooth M/S (EDF + CBS).

802.15.4 network, the proposed approach produces better performance. Periodic TH/WL percentage measured is 82.5% against 74.8% measured in a Bluetooth scenario.

Figure 15 shows deadline miss ratio results in IEEE 802.15.4 and Bluetooth, respectively. In an IEEE 802.15.4 network, we measured less deadline miss percentage than Bluetooth for both periodic (4.1% versus 6.4%) and aperiodic traffic flows (1.9% versus 2.1%).

Figures 16 and 17 finally show average delay for periodic and aperiodic packets, respectively. It is possible to see that Bluetooth produces on average a delay of 30.26 ms lower than IEEE 802.15.4 in case of periodic traffic flows while, in case of aperiodic traffic flows, the estimated delay is on average 19.10 ms lower using a Bluetooth network.

5. Conclusions

In this paper, a real-time scheduling algorithm, for both periodic and aperiodic traffic management, applied to networks based on IEEE 802.15.4 and Bluetooth, respectively, has been presented. The main aim of this work is to reduce, as much as possible, the packet loss on the wireless channel increasing, at the same time, the reliability of the wireless technology in terms of TH/WL ratio and average delay. Furthermore, a comparison between the IEEE 802.15.4 and Bluetooth solutions has been conducted in order to identify the more suitable communication protocol for industrial process control systems. As shown, our approach improves performance of both IEEE 802.15.4 and Bluetooth standards. Moreover, the simulation campaign clearly demonstrates that the combinated use of EDF (for periodic traffic flows management) and CBS (per aperiodic traffic flows management) determines more enhancements in IEEE 802.15.4 than Bluetooth. In fact, the TH/WL percentage of periodic flows measured in a IEEE 802.15.4 network is 82.5% against 74.8% measured in a Bluetooth scenario. Similarly, better performance has been obtained in terms of deadline miss ratio for both periodic (4.1% versus 6.4%) and aperiodic (1.9% versus 2.1%) flows. Instead, as previously shown in Section 4, in terms of average delay measured, our approach produces better performance using Bluetooth than IEEE 802.15.4. The measured delay of

periodic traffic flows in Bluetooth is 30.26 ms lower than IEEE 802.15.4 while, in case of aperiodic traffic flows, the estimated delay in Bluetooth is on average 19.10 ms lower than IEEE 802.15.4. In conclusion, in applications where the main requirement is to have low packet loss reducing, at the same time, the deadline miss, it is preferable to use IEEE 802.15.4. On the contrary, in applications where it is important to ensure low network communication latencies, it is preferable to use Bluetooth.

References

[1] T. Naumowicz, R. Freeman, H. Kirk et al., "Wireless sensor network for habitat monitoring on Skomer Island," in *Proceedings of the 35th Annual IEEE Conference on Local Computer Networks (LCN '10)*, pp. 882–889, Denver, Colo, USA, October 2010.

[2] M. Collotta, V. Conti, G. Pau, G. Scatà, and S. Vitabile, "Fuzzy techniques for access and data management in home automation environments," *Journal of Mobile Multimedia*, vol. 8, no. 3, pp. 181–203, 2012.

[3] M. Collotta, G. Scatà, and G. Pau, "A priority-based CSMA/CA mechanism to support deadline-aware scheduling in home automation applications using IEEE 802. 15. 4," *International Journal of Distributed Sensor Networks*, vol. 2013, Article ID 139804, 12 pages, 2013.

[4] M. Collotta, G. Nicolosi, E. Toscano, and O. Mirabella, "A ZigBee-based network for home heating control," in *Proceedings of the 34th Annual Conference of the IEEE Industrial Electronics Society (IECON '08)*, pp. 2724–2729, Orlando, Fla, USA, November 2008.

[5] J. Byun, B. Jeon, J. Noh, Y. Kim, and S. Park, "An intelligent self-adjusting sensor for smart home services based on ZigBee communications," *IEEE Transactions on Consumer Electronics*, vol. 58, no. 3, pp. 794–802, 2012.

[6] M. Collotta, G. Pau, V. M. Salerno, and G. Scatà, "A novel road monitoring approach using wireless sensor networks," in *Proceedings of the 6th International Conference on Complex, Intelligent and Software Intensive Systems (CISIS '12)*, vol. 2012, pp. 376–381, 2012.

[7] A. Pascale, M. Nicoli, F. Deflorio, B. Dalla Chiara, and U. Spagnolini, "Wireless sensor networks for traffic management and road safety," *IET Intelligent Transport Systems*, vol. 6, no. 1, pp. 67–77, 2012.

[8] M. Collotta, L. Gentile, G. Pau, and G. Scatà, "A dynamic algorithm to improve industrial wireless sensor networks management," in *Proceedings of the 38th Annual Conference of IEEE Industrial Electronics (IECON '12)*, pp. 2802–2807, Montreal, Canada, 2012.

[9] M. Collotta, L. L. Bello, and E. Toscano, "A proposal towards flexible wireless communication in factory automation based on the IEEE 802.15.4 protocol," in *Proceedings of the IEEE Conference on Emerging Technologies and Factory Automation (ETFA '09)*, September 2009.

[10] J. Lin and C. Liu, "A monitoring system based on wireless sensor network and an SoC platform in precision agriculture," in *Proceedings of the 11th IEEE International Conference on Communication Technology (ICCT '08)*, pp. 101–104, Hangzhou, China, November 2008.

[11] N. P. Karthickraja, V. Sumathy, and J. Ahamed, "A novel hybrid routing protocol for data aggregation in agricultural applications," in *Proceedings of the IEEE International Conference on Communication Control and Computing Technologies (ICCCCT '10)*, pp. 227–231, Ramanathapuram, India, October 2010.

[12] Y. Chen, W. Shen, H. Huo, and Y. Xu, "A smart gateway for health care system using Wireless Sensor Network," in *Proceedings of the 4th International Conference on Sensor Technologies and Applications (SENSORCOMM '10)*, pp. 545–550, Venice, Italy, July 2010.

[13] F. Pu, C. Li, T. Gao, J. Pan, and J. Li, "Design and implementation of a wireless sensor network for health monitoring," in *Proceedings of the 4th International Conference on Bioinformatics and Biomedical Engineering (iCBBE '10)*, pp. 1–4, Chengdu, China, June 2010.

[14] M. Collotta, L. L. Bello, E. Toscano, and O. Mirabella, "Dynamic load balancing techniques for flexible wireless industrial networks," in *Proceedings of the 36th Annual Conference of the IEEE Industrial Electronics Society (IECON '10)*, pp. 1329–1334, Glendale, Calif, USA, November 2010.

[15] M. Collotta, G. Pau, V. M. Salerno, and G. Scatà, "A distributed load balancing approach for industrial IEEE 802. 11 wireless networks," in *Proceedings of the IEEE Conference on Emerging Technologies & Factory Automation (ETFA '12)*, pp. 1–4, 2012.

[16] M. Collotta, G. Pau, V. M. Salerno, and G. Scata, "A fuzzy based algorithm to manage power consumption in industrial Wireless Sensor Networks," in *Proceedings of the 9th IEEE International Conference on Industrial Informatics (INDIN '11)*, pp. 151–156, Caparica, Portugal, July 2011.

[17] "802. 15. 4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (LR- WPANs)," June 2006 IEEE standard for information technology. Part 15. 4.

[18] Bluetooth Specification Version 3. 0 + HS, Bluetooth SIG, 2009.

[19] G. Mulligan, "The 6LoWPAN architecture," in *Proceedings of the 4th Workshop on Embedded Networked Sensors (EmNets '07)*, pp. 78–82, ACM, June 2007.

[20] J. Song, S. Han, A. K. Mok et al., "WirelessHART: applying wireless technology in real-time industrial process control," in *Proceedings of the 14th IEEE Real-Time and Embedded Technology and Applications Symposium (RTAS '08)*, pp. 377–386, St. Louis, Mo, USA, April 2008.

[21] J. L. Sevillano, D. Cascado, F. Díaz Del Río, S. Vicente, G. Jimenez, and A. Civit-Balcells, "Soft real-time communications over Bluetooth under interferences from ISM devices," *International Journal of Communication Systems*, vol. 19, no. 10, pp. 1103–1116, 2006.

[22] M. Collotta, L. Lo Bello, and O. Mirabella, "Deadline-aware scheduling policies for Bluetooth networks in industrial communications," in *Proceedings of the IEEE 2nd International Symposium on Industrial Embedded Systems (SIES '07)*, pp. 156–163, Lisbon, Portugal, July 2007.

[23] M. Collotta, L. L. Bello, and O. Mirabella, "An innovative frequency hopping management mechanism for Bluetooth-based industrial networks," in *Proceedings of the 5th International Symposium on Industrial Embedded Systems (SIES '10)*, pp. 45–50, Trento, Italy, July 2010.

[24] A. Koubâa, M. Alves, E. Tovar, and A. Cunha, "An implicit GTS allocation mechanism in IEEE 802.15.4 for time-sensitive wireless sensor networks: theory and practice," *Real-Time Systems*, vol. 39, no. 1–3, pp. 169–204, 2008.

[25] A. Koubaa, M. Alves, and E. Tovar, "GTS allocation analysis in IEEE 802. 15. 4 for real-time wireless sensor networks," in *Proceedings of the 20th International Parallel and Distributed Processing Symposium (IPDPS '06)*, 2006.

[26] J. Mišic and V. B. Mišic, "Duty cycle management in sensor networks based on 802. 15. 4 beacon enabled MAC," *Ad Hoc and Sensor Wireless Networks Journal*, vol. 1, no. 3, pp. 207–233, 2005.

[27] J. Mišić, S. Shafi, and V. B. Mišić, "The impact of MAC parameters on the performance of 802.15.4 PAN," *Ad Hoc Networks*, vol. 3, no. 5, pp. 509–528, 2005.

[28] Y. Xue, S. L. Ho, M. Yang, P. Kumarawadu, H. H. Ghenniwa, and W. Shen, "Performance evaluation of NS-2 simulator for wireless sensor networks," in *Proceedings of the Canadian Conference on Electrical and Computer Engineering (CCECD '07)*, pp. 1372–1375, Vancouver, Canada, April 2007.

[29] J. A. Stankovic, T. F. Abdelzaher, C. Lu, L. Sha, and J. C. Hou, "Real-time communication and coordination in embedded sensor networks," *Proceedings of the IEEE*, vol. 91, no. 7, pp. 1002–1022, 2003.

[30] V. Rajendran, K. Obraczka, and J. J. Garcia-Luna-Aceves, "Energy-efficient, collision-free medium access control for wireless sensor networks," *Wireless Networks*, vol. 12, no. 1, pp. 63–78, 2006.

[31] E. Egea-López, J. Vales-Alonso, A. S. Martínez-Sala, J. García-Haro, P. Pavón-Mariño, and M. V. Bueno Delgado, "A wireless sensor networks MAC protocol for real-time applications," *Personal and Ubiquitous Computing*, vol. 12, no. 2, pp. 111–122, 2008.

[32] W. Zhang, H. Zhu, and G. Cao, "On improving the performance of bluetooth networks through dynamic role management," Tech. Rep. CSE-01-018, Pennsylvania State University.

[33] S. Abhyankar, R. Toshiwal, C. Cordeiro, and D. Agrawal, "On the application of traffic engineering over bluetooth Ad Hoc networks," in *Proceedings of the 6th ACM International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM '03)*, pp. 116–123, September 2003.

[34] L. L. Bello, N. Torrisi, and O. Mirabella, "New operating modes for bluetooth networks in distributed process control systems," in *Proceedings of the 29th Annual IEEE International Conference on Local Computer Networks (LCN '04)*, pp. 94–101, November 2004.

[35] L. Lo Bello, M. Collotta, O. Mirabella, and T. Nolte, "Approaches to support Real-Time traffic over Bluetooth," in *Proceedings of RTN 05, the 4st International Workshop on Real-Time Networks, in Conjunction with the 17th Euromicro Conference on Real-Time Systems (ECTRS '05)*, pp. 47–50, 2005.

[36] G. C. Buttazzo, *Hard Real-Time Computing Systems—Predictable Scheduling Algorithms and Applications*, Springer, 3rd edition, 2011.

[37] R. Devillers and J. Goossens, "Liu and Layland's schedulability test revisited," *Information Processing Letters*, vol. 73, no. 5, pp. 157–161, 2000.

[38] http://www.omnetpp.org/.

[39] http://www.cs.uc.edu/~cdmc/ucbt/.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 591809, 8 pages
http://dx.doi.org/10.1155/2013/591809

Research Article

Data Storage Scheme Supporting for Multidimensional Query

Keji Mao, Xiaomin Zhao, Qike Shao, Wenxiu He, Yanqiang Ou, and Qingzhang Chen

College of Computer Science and Technology, Zhejiang University of Technology, Zhejiang 310023, China

Correspondence should be addressed to Keji Mao; maokeji@zjut.edu.cn

Received 20 December 2012; Revised 17 March 2013; Accepted 7 April 2013

Academic Editor: Tai-hoon Kim

The query and storage of data is very important in wireless sensor networks (WSN). It is mainly used to solve how to effectively manage the distributed data in the monitored area with extremely limited resources. Recent advances in technology have made the number of the sensing modules in one sensor develop from single to multiple. The existing storage scheme for one-dimensional data is not suitable for the multidimensional data or costs too much energy. We proposed a kind of data storage scheme supporting multidimensional query inspired by K-D tree. The scheme can effectively store the high-dimensional similar data to the same piece of two-dimensional area. It can quickly fix the storage area of the event by analyzing the query condition and then fetch back the query result. Meanwhile the scheme has a certain degree of robustness to packet loss and node failure. Finally the experiment on the platform of Matlab showed that our scheme has some advantages compared with the existing methods.

1. Introduction

A wireless sensor network is composed of a large number of sensor nodes deployed in the monitored area and these nodes communicate through wireless waves. The basic function of WSN is accessing the information. The user can get the information of objects through these nodes and can even control some of them [1].

As the basic component of WSN, a sensor node usually consists of several modules, just like sensing module, processing module, communication module, and power module [2]. Earlier sensor node is simple and usually includes only one sensor, such as temperature sensor or humidity sensor. It gets only one numerical value when it detected events and these events are called one-dimensional events. The queries for these events are called one-dimensional query. Recent advances in technology have made the number of sensor of the sensing model develops from single sensor to multiple sensors and the sensing model can get several numerical values at the same time, such as temperature, humidity, light, and pressure. When it detects an event, it gets a value like ⟨temperature, humidity, light, pressure $\cdots$⟩ and these events are called multidimensional events. According to the number of dimensions and the range size of query, multidimensional query can be divided into multidimensional exact match point query, multidimensional partial match point query, multidimensional exact match range query, and multidimensional partial match range query [3].

Multidimensional events emerge with the development of technology. The existing data storage schemes are designed for the one-dimensional event and these schemes cannot be used for multidimensional event, or it costs too much energy in multidimensional event query. Limited energy is an important feature of WSN and the life time of the whole network is an important indicator of its performance; meanwhile the long-time stable running is also a significant security for the usage in reality. So how to design a data storage scheme to support the multidimensional query, which can avoid amount of useless queries, reduce the response time, save the energy of the nodes, and extend the life of the network, is faced by all scholars.

The existing schemes supporting for multidimensional query include DIM [4] and Pool [5]. DIM is the first algorithm used to solve the problem. It stores the event by using the code of the event and resolves the query into many subqueries and gets the result, respectively. Pool addresses some weakness of DIM such as hot spot and scalability problems and further proposes a storage scheme to group index nodes together as a data pool to preserve the neighborhood property of nearby multidimensional data; however, their group mechanism is only based on the greatest and the second greatest attribute values, which has not been proved to be efficient

enough for high-dimensional data processing. The literature [6] proposed a secure and reliable data distribution scheme in unattended WSNs. Work [7] proposed an energy-efficient query algorithm to query the top-k value of sensed data. MDCS is a multiresolution data compression and storage strategy [8]. Work [9] proposed a bouncing-track-based data storage and discovery scheme in WSN. The literature [10] proposed a secure and reliable data distribution scheme in WSN.

The existing storage schemes supporting multidimensional query mainly give solutions from the theoretical point, but there are still many shortcomings. Hot spots: the storage node and the nodes nearby may quickly become the hot spots because of excessive communication. Scalability problem: as the number of sensor nodes in the system increases, the amount of message exchanges among sensors increases too fast. Query costs too much: the cost is too high in the present schemes. Hard to implement: most of the existing schemes require the accurate division of the network to ensure a grid only has only one or a few nodes and this is hard to carry out in real [11].

Faced with the emergence of the multidimensional events and flaws of the existing data storage schemes, we propose a new method inspired by K-D tree. It can solve the flaws in some degree, effectively reduce the energy cost, and extend the life of the network.

2. Design of Data Storage Scheme Supporting for Multidimensional Query

2.1. K-D Tree Introduction. K-D tree [12] is a data structure that can effectively store multidimensional data. K represents the dimensionality of the search space; K-D tree is evolved from the binary tree. The difference is that each layer represents a dimension and you can use the value of each dimension as a splitter to make decisions. The top layer node is chosen from any of the k dimensions. The second layer is divided by another dimension. When the value of depth of the tree is larger than the value of dimensions, you can reuse these dimensions.

The advantage of the K-D tree is that the data are classified according to the value of each dimension. The closer the value of each dimensional data the more probable the data may be classified together, this idea can be directly applied to multidimensional queries in WSN. In the progress of designing the data storage supporting multidimensional querying, first, you should map the data in high-dimensional space to two-dimensional space and maintain the local nature [13]. Second, you can reuse the dimension to split the data into smaller range. So you can give a more accurate position in querying. Third, you can split the area according to each value, and this value also can be applied to the division of monitoring area in WSN.

The basic idea of data storage is as follows: suppose the dimension of monitored event is k; first, we construct a binary tree of depth of d, $d \geqslant k$. Each layer of the tree is corresponded to each dimension of the event. When $d \leqslant k$, each layer of the tree is divided according to 0.5. When $k \leqslant d \leqslant 2k$, 0.25 and 0.75 are used as the splitter values. At last all data are stored in leaf nodes.

2.2. Data Storage Mechanism

2.2.1. Mapping an Event to a Storage Area. Suppose the wireless sensor nodes are uniformly distributed in the monitoring area, and nodes can communicate through the exchange of information. All nodes know their coordinates in the area. When the node detects an event, it will get a set of data elements about the event. We use E to represent the k-dimensional event: $E = \langle V_1, V_2, \ldots, V_k \rangle$, in which each attribute value has normalized. For the query issued by the user, we use Q to represent $Q = \langle [L_1, U_1], [L_2, U_2], \ldots, [L_h, U_h] \rangle$, $h \leq k$; in which L_i and U_i, respectively, represent the minimum and maximum attributes of i.

First, we construct a virtual binary tree before mapping the event to the storage area, then we divide the storage area and insert event on the basis of the binary tree. The key point is how to fix the depth of the tree, because it will decide the number of storage area. Assume the depth of the tree is d and the dimension of event is k. To each dimension of events for a comparison, the depth needs to fix the condition: $d \geqslant k$. For the maximum value about d, we assume nodes in each partitioned grid will receive the data in all directions and have the same probability. In order to store data with the best balanced manner, we set a storage node in each direction, so the grid should have at least 8 nodes. Assume the number of nodes in the network is N, so the number of grid is $N/8$ at most and the maximum depth of the tree is $d \leq \sqrt{N/8}$. Therefore, the depth of the tree should meet the condition: $k \leq d \leq \sqrt{N/8}$. The value of d can be flexibly selected according to the size of monitoring area.

When the depth of the tree d is determined, the network monitoring area is divided into 2^d storage area. When the event is generated in the network, it should be stored into appropriate storage area. First, we should encode the event. For event $E = \langle V_1, V_2, \ldots, V_k \rangle$, we assign a d-bit binary string. For every bit i, i changes from 1 to k; if $0 \leq V_i < 0.5$, this bit is encoded as 0; if $0.5 \leq V_i \leq 1$, this bit is encoded as 1. When i changes from $k+1$ to $2k$, if $(0 \leq V_{i-k} < 0.25) \cup (0.5 \leq V_{i-k} < 0.75)$, this bit is encoded as 0; if $(0.25 \leq V_{i-k} < 0.5) \cup (0.75 \leq V_{i-k} \leq 1)$, this bit is encoded as 1. This work stops until the d-bit numbers are all assigned. At last, the event E is encoded into E_{code}.

In E_{code}, the odd bit divides the area symmetry in left and right. Left is 0 and right is 1. The event bit divides the area into above and below. Above is 0 and below is 1. So each binary string at last determines a storage area. This storage area is defined by the lower left coordinates $(X_{\text{low}}, Y_{\text{low}})$ and upper right coordinates $(X_{\text{up}}, Y_{\text{up}})$.

Suppose $E_{\text{code}} = [x_1, y_1, x_2, y_2, \ldots, x_{d/2}, y_{d/2}]$; we first calculate the lower bound X_{low} on the x-axis as follows:

$$X_{\text{low}} = x_1 (0.5) + x_2 (0.5)^2 + \cdots + x_{d/2} (0.5)^{d/2}. \quad (1)$$

We get the lower bound of y-axis as follows:

$$Y_{\text{low}} = y_1 (0.5) + y_2 (0.5)^2 + \cdots + y_{d/2} (0.5)^{d/2}. \quad (2)$$

Therefore we define a matrix M as follows:

$$M = \begin{bmatrix} \left(\frac{1}{2}\right) & 0 \\ 0 & \left(\frac{1}{2}\right) \\ \left(\frac{1}{2}\right)^2 & 0 \\ 0 & \left(\frac{1}{2}\right)^2 \\ \left(\frac{1}{2}\right)^{d/2} & 0 \\ 0 & \left(\frac{1}{2}\right)^{d/2} \end{bmatrix}. \quad (3)$$

We can calculate R_{low} as follow:

$$R_{\text{low}} = [X_{\text{low}}, Y_{\text{low}}] = E_{\text{code}} \cdot M. \quad (4)$$

After getting the lower bound of the storage area, we need to find the upper bound X_{up} and Y_{up}. We first make XOR operation on E_{code}:

$$\overline{E_{\text{code}}} = E_{\text{code}} \oplus J_d \quad J_{d/2} = [1, 1, \ldots, 1]_{1\times d}. \quad (5)$$

Then we calculate $\overline{R_{\text{up}}}$:

$$\overline{R_{\text{up}}} = \left[\overline{X_{\text{up}}}, \overline{Y_{\text{up}}}\right] = \overline{E_{\text{code}}} \cdot M. \quad (6)$$

At last we calculate the real upper bound R_{up}:

$$R_{\text{up}} = \left[X_{\text{up}}, Y_{\text{up}}\right] = J_2 - \overline{R_{\text{up}}}. \quad (7)$$

Finally the storage area is defined by R_{low} and R_{up}: $R = [X_{\text{low}} \sim X_{\text{up}}, Y_{\text{low}} \sim Y_{\text{up}}]$.

2.2.2. Routing an Event to a Storage Area. After getting the storage area of an event, we need to route the event to this area to store. The progress is as follows: the node that detected the event first compares its geographic coordinate with the coordinate of storage area. If the node's coordinate falls in the range of storage area's coordinate, the node will store the event locally. Otherwise, the node will attempt to route the event to the corresponding storage area with GPSR (GPSR: Greedy Perimeter Stateless Routing for Wireless Networks).

GPSR is a geographic routing protocol, the benefit of which is that it can route the packet to the destination with greedy manner after known coordinate of the source node and destination node [14]. In addition to the protocol's own information, the package also carried the coordinate of the storage area. The packet format is as follows:

ID $\cdots$ X_{low} Y_{low} X_{up} Y_{up} X_{mid} Y_{mid}.

X_{mid} and Y_{mid} represent the coordinate of the center of R. They are the destination address of GPSR. In the routing progress, every node will compare its own coordinate with them. When a node meets the condition of $(X_{\text{low}} \le x \le X_{\text{up}}) \cap (Y_{\text{low}} \le y \le Y_{\text{up}})$, that means the event has been routed to the storage area. The next step is selecting a node to store the event. In order to ensure load balancing, the selection of storage nodes needs to meet certain conditions.

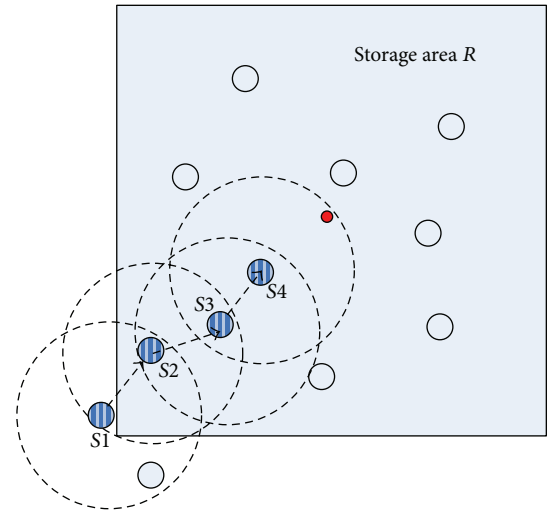


Figure 1: Load balance in storage area.

2.2.3. Load Balance in the Storage Area. After the event has been routed to the storage area, we need to select a node to store the event. There are many methods. In this scheme we operate like this: we select the node along the direction of the event coming from and closest to the center of the area to store as Figure 1 shows. Suppose the lower left coordinate of R is $R_{\text{low}}(X_{\text{low}}, Y_{\text{low}})$, the upper right coordinate of R is $R_{\text{up}}(X_{\text{up}}, Y_{\text{up}})$, and the middle coordinate of R is $R_{\text{mid}}(X_{\text{mid}}, Y_{\text{mid}})$; then the storage node $S(x, y)$ should meet the following condition: Min $\sqrt{(x - X_{\text{mid}})^2 + (y - Y_{\text{mid}})^2}$ and $(X_{\text{low}} \le x \le X_{\text{up}}) \cap (Y_{\text{low}} \le y \le Y_{\text{up}})$.

The storage capacity of every node in the network is limited. We set a threshold to node's memory load. When the capacity of the node reaches the threshold, the node cannot receive events. In the progress of selecting storage nodes, the node will be skipped. By using this method, every node will have the opportunity to store events.

2.3. Data Query Mechanism. The query processing mechanism contains two parts: query resolving and retrieval route. The design of the query resolving is determining a set of relevant range spaces which can provide answers to the query. The retrieval route is fetching match events to the query from set space.

2.3.1. Query Resolving. Multidimensional queries include point queries and range queries. Point querying is identical to inserting an event, but range querying includes multidimensional exact match range querying and multidimensional partial match range querying. It is related to more storage areas, so you need to resolve the query conditions first. The following parts, respectively, described their solutions.

(1) Query Resolving Mechanism for Exact Match Range Queries. We represent the event in the network with $E = \langle V_1, V_2, \ldots, V_k\rangle$, when the user sends a query like $Q = \langle [L_1, U_1], [L_2, U_2], \ldots, [L_h, U_h]\rangle$ and it satisfies $(h = k) \cap (L_i \le U_i)$, $1 \le i \le k$. We call the query as exact match range querying. How

to split the query is as follows: if the depth of the binary tree is identical to the number of dimensions, then we check Q's ith attribute whether contains the value 0.5. If it contains, then we divide the Q in two subqueries. If the depth is larger, then the query will be divided into smaller ranges. For example, if you send a query $Q = \langle[0.3, 0.6], [0.7, 0.8], [0.3, 0.4]\rangle$ and the depth of the binary tree is 4, you will get four subqueries after resolving:

$$\begin{aligned} Q_1 &= \langle[0.3, 0.5), [0.7, 0.8], [0.3, 0.4], [0.3, 0.5]\rangle, \\ Q_2 &= \langle[0.3, 0.5), [0.7, 0.8], [0.3, 0.4], [0.5, 0.6]\rangle, \\ Q_3 &= \langle[0.5, 0.6], [0.7, 0.8], [0.3, 0.4], [0.3, 0.5]\rangle, \\ Q_4 &= \langle[0.5, 0.6], [0.7, 0.8], [0.3, 0.4], [0.5, 0.6]\rangle. \end{aligned} \tag{8}$$

The binary strings of the four subqueries are 0101, 0100, 1101, and 1100 and they, respectively, correspond with four storage areas in the network. 0101 and 0100, 1101 and 1100 are different with the last bit. We can certify their storage areas are adjacent. So if sub-queries are just different with the last bit, we can combine the two sub-queries into one query and get rid of the last bit. In this example, the combined binary strings are 101 and 110. The benefit of the mechanism is reducing the number of routing path from four to two to save the power energy.

(2) Query Resolving Mechanism for Partial Match Range Queries. Assume that $Q = \langle[L_1, U_1], [L_2, U_2]\rangle = \langle[0.3, 0.6], [0.7, 0.8]\rangle$ is a three-dimensional partial match range query. The value of the three-dimensional can be anything. That means there are more results satisfing the condition and the searching area for the query will be extended. For this kind of queries, we can reset the range of each unspecified attribute as $[0, 1]$, then the query became $Q = \langle[0.3, 0.6], [0.7, 0.8], [0, 1]\rangle$, then we can use the resolving method of exact match range queries.

2.3.2. Retrieval Route. Routing queries is similar to inserting an event. You can get several sub-queries after resolving and combining these queries. And every binary string of the sub-queries corresponds to a storage area. By using GPSR routing protocol, these sub-queries will route to their corresponding storage areas. Upon receiving the sub-query, a node would flood the sub-query to nodes in the same space. The nodes answer the sub-query with matching events values. The retrieved qualifying events are sent back to the user through the same path.

3. Simulation and Analysis

3.1. Experiment Parameters and Performance Evaluation Factors. We make the simulation on the platform of Matlab and compare our scheme to DIM and Pool. Because the exact point query is simple, it is hard to show the differences of each algorithm. We choose range query as the comparison object. The experiment parameters are the same as those in DIM and Pool.

(1) Sensor Nodes Setting. In the experiment, we uniformly place sensor nodes in the entire field. The number of sensor nodes varies from 300 to 1800. Each node has a radio range of 40 m and has on average 20 nodes within its nominal radio range. Every node's storage capacity is 20 events.

(2) Events Setting. In the experiment, each sensor node on average generates three events and each event has four dimensions. If there are N nodes in the network, the total events are $3N$.

(3) Exact Match Query Setting. We employ the four query size distributions which are also used in DIM and Pool to emulate the range sizes specified in a query for the fair comparison. They are uniform in width, bounded width, algebraic width, and exponential width. Here we briefly describe the concepts of these size distributions. Two parameters mp and rs are used to generate attribute value range of a query. mp stands for the midpoint of an attribute value range in a query and rs stands for its range size. When mp and rs of a value range are given, the range can be represented as [MAX (mp − rs/2, 0), MIN (mp + rs/2, 1)]. We uniformly select mp with $[0, 1]$ for all queries; rs on the other hand is different in different size distribution.

For a uniform range size distribution, the range size for each dimension is uniformly distributed in $[0, 1]$. To generate such queries, we uniformly select rs within the range $[0, 1]$. For the bounded uniform distribution the range size on each dimension of a query is uniformly distributed in a predefined B. In all the experiments, the bound is set to be 0.5. Thus, to generate queries with bounded uniform size distribution, rs is uniformly selected within $[0, 0.5]$. In an algebraic distribution, most queries have a small range on each dimension. rs is more likely to be selected within $[0, 0.25]$. Finally, for the exponential distribution, all queries have a small range size on all the dimensions. More specifically, the range size on each dimension is uniformly distributed over $[0, 0.25]$.

(4) Partial Match Query Setting. A partial match query occurs when one or more attribute value ranges are unspecified. In that case, we will consider those unspecified ranges as the entire value range $[0, 1]$. For convenience in discussion, we refer to an m-partial match query as a query with m specified value ranges. For a four-dimensional partial match query, m can be 1 or 2 or 3. When $m = 3$, that represents the value of one dimension is not specified. For a four-dimensional query, that could be any one. If the unspecified value range is nth attribute in the m-dimensional query, we can use $m@n$ to represent it. The reason of having the distinction of the partial match query is that the performance is strongly dependent on which attribute in the query has an unspecified value range. It performs quite differently if the unspecified range occurs at different attribute.

The performance metrics employed in the experiments are the same as those used in DIM and Pool. They are average

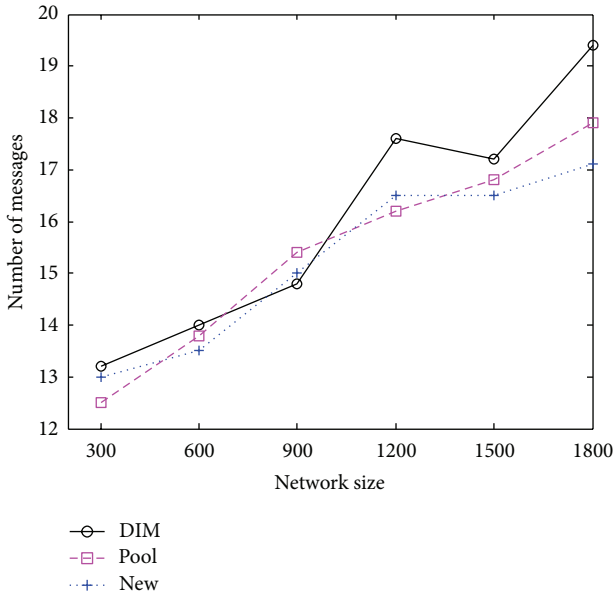


Figure 2: Data insertion cost.

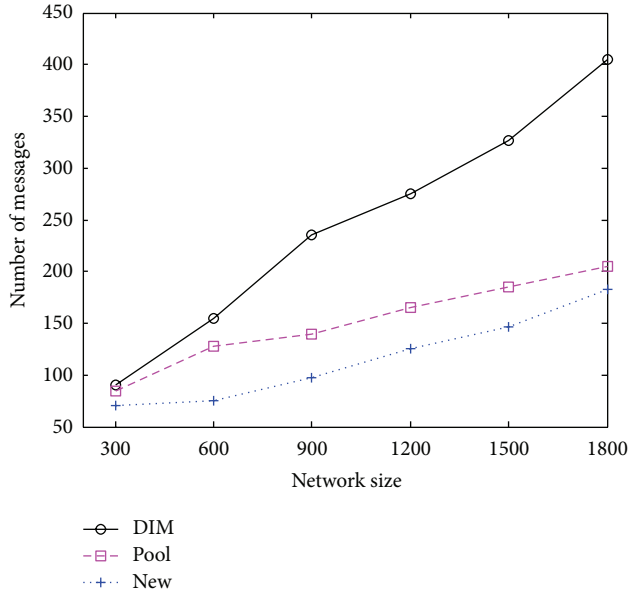


Figure 3: Uniform range size distributions.

insertion cost, average query delivery cost, and the number of hot spot.

Average Insertion Cost. It measures the average number of messages required to insert an event into the network.

Average Query Delivery Cost. It measures the average number of messages required to route a query message to all the relevant nodes in the network.

Number of Hot Spot. In order to achieve load-balancing effect, we set a threshold to every node. Once the number of stored events reaches the threshold, it changed into a host spot.

3.2. Analysis of Simulation

3.2.1. Average Insertion Cost. Configured with the above parameters, we investigated the average insertion cost under different size of network. The result is as shown in Figure 2.

We can find that there is no difference between these three approaches. This is not surprising since the insertion of an event is simply routing a data packet from one sensor node to another, and all the three approaches use the GPSR. We can also find that Pool and our algorithm are a slightly better than the DIM, as in the routing progress there is no clear coordinate of the destination node. It routes the data package to the destination step by step by comparison of the coding, which sent more packages than the other two methods.

3.2.2. Average Query Delivery Cost

(1) Comparisons on Exact Match Queries. We compared the three algorithms under four different range size distributions and the results are as shown in Figures 3, 4, 5, and 6.

We can observe some interesting facts from the four figures.

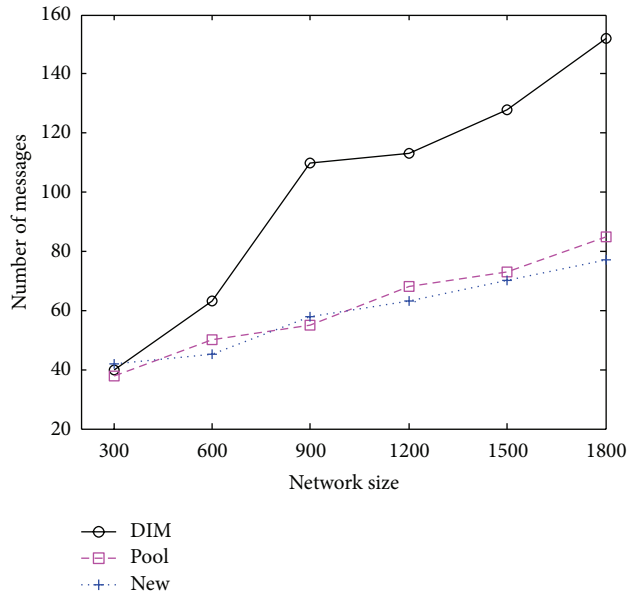


Figure 4: Bounded Uniform query size distributions.

First, all the figures reveal that the range size of a query strongly affects the performance of the three approaches. For instance, all the approaches cost much less in exponential size distribution than in uniform size distribution. This is mainly because data are stored according to splitters like 0.5 and 0.25. The smaller the range size is, the easier a query gets the result. A query therefore has to visit a large amount of storage areas in order to acquire all the answers in uniform size distribution.

Second, the performance result reveals that our algorithm is scalable than Pool, and Pool is scalable than DIM. In DIM all sensors are treated as index nodes. The network needs to construct a huge structure of tree first and that definitely leads

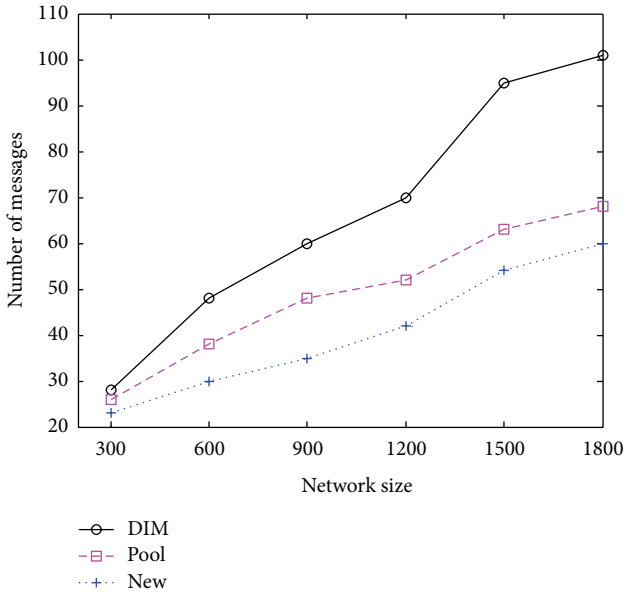


Figure 5: Algebraic query size distributions.

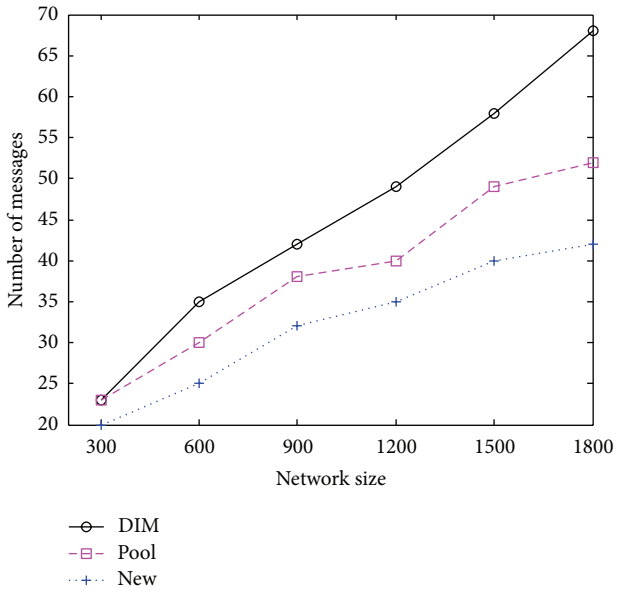


Figure 6: Exponential range size distributions.

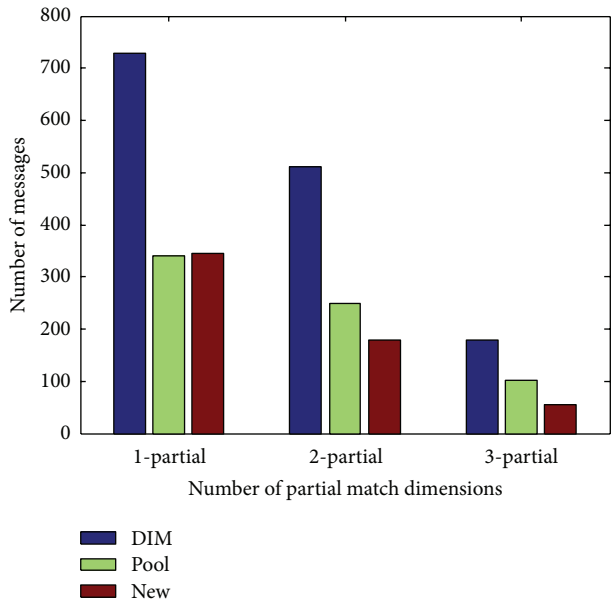


Figure 7: Comparisons on partial match dimensions.

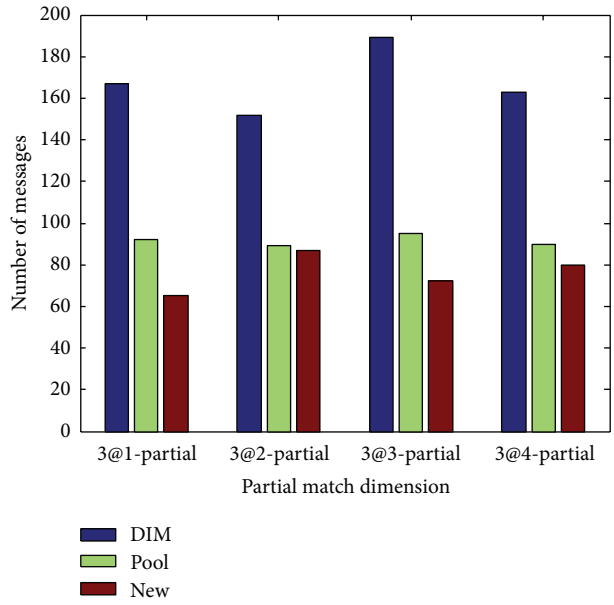


Figure 8: Partial match dimension order.

to the flood of messages. In Pool the number of pool will not change when the size of network turned larger. A package needs to visit more nodes to get to the destination. In our method the size of the tree is adjustable. We can control the size of flooding area by changing the depth of the tree.

(2) Comparisons on Partial Match Queries. We fixed the size of the network as 900 sensor nodes and compared these methods under one, two, and three of dimensions with specified ranges. The results are as shown in Figure 7 and Figure 8.

The result in Figure 7 shows that the more dimensions in a query with an unspecified range, the higher the query processing cost would be incurred. The reason is that the more unspecified ranges are, the more index nodes would be accessed which incurs a higher cost. In addition, in the same dimension with unspecified range, DIM costs much more, because it randomly sends the message and the message may route to some unrelated nodes. In one-dimensional partial query, the effect of our method is the same as Pool. The reason is the flooding area becomes large when the number of unspecified range is high. But our method performs better than Pool when the number of unspecified range changed small. The reason is Pool is always searching in the fixed area but ours has changed.

In partial match queries, the location of unspecified dimensions can be changed when its amount is fixed. Figure 8 shows the result of a three-dimensional query.

We can see from Figure 8 that the change of the location of unspecified dimension does not have too much influence to the three methods and Pool is the most stable one. In Pool, the storage area is stable and the number is fixed. The change of the location of unspecified dimension just influences the length from sink to the storage area, but the change in DIM and our scheme can affect the amount and position of the storage area. We can see in Figure 8 that the number of messages is a fluctuating state.

3.2.3. Comparisons on Hot Spot. A good algorithm should have a good mechanism to load balancing. All of the three algorithms have this kind of mechanism. We fixed the size of the network as 900 sensor nodes. Every node sends three packages and the threshold is 20 packages. The result is shown in Figure 9.

We can find that DIM performs worse than Pool, and Pool performs worse than our scheme. In scheme, every node chose a backup node. When it reaches the threshold, the backup node starts to replace it and store data. An event is corresponding to a range and a range is corresponding to a node. So an event directly figures a storage node. In this case, the storage node will soon reach the threshold. In Pool, the storage area is determined by both the events' greatest and the second greatest attribute values. This in some extent divides the similar high-dimensional events again and ensures the storage nodes do not grow too fast. In our scheme, storage area relies on the binary string of the events. The storage node is chosen from the direction of event and it is closest to the center of the storage area. The similar high-dimensional events coming from different directions may store at different nodes. This is the mechanism to load balancing in our scheme. Figure 9 also shows that this method can indeed slow down the production of the number of hot spots.

4. Summary

Data storage and query is an important aspect of data management in WSN. A good storage scheme can give the convenience for the query, save the energy for the network, and greatly extend the network lifetime. For the emergence of multidimensional events, scholars have proposed some ways to ensure the effectiveness of energy. The data storage scheme in this paper is inspired by K-D tree. It first calculates the storage area by encoding the events, then it routes the package to there. The selection of the storage node is related to the direction of the event. In a storage area, there can be different storage nodes and this in some extent ensures the load balancing in the network. The scheme supports multidimensional query through resolving, encoding, and combing the query. In addition, for the integrity of the algorithm the scheme has robustness to package loss and node failure. Finally, we made experiment on the platform of Matlab. The result shows the scheme has much advantage than DIM and is slightly better than Pool, but the complexity and robustness of the scheme are better than both of them.

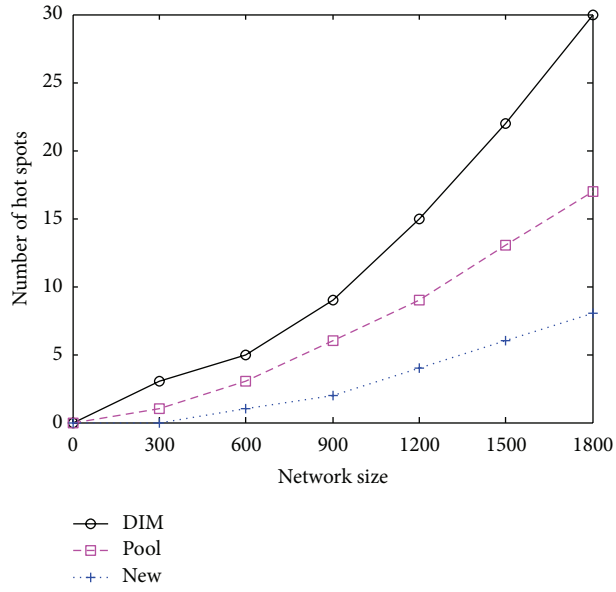


Figure 9: Comparisons of hot spot.

Acknowledgments

This work is supported by the National Natural Science Foundation of China under Grant no: 61104095, the Natural science Foundation of Zhejiang Province under Grant nos: LY12F02036; Y1110915, the Science and Technology Research Project of Zhejiang Province under Grant nos: 2011C21014; 2012C33085; 2012C33073.

References

[1] D. Culler, D. Estrin, and M. Srivastava, "Overview of sensor networks," *Computer*, vol. 37, no. 8, pp. 41–49, 2004.

[2] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–114, 2002.

[3] R. Gummadi, X. Li, R. Govindan, C. Shahabi, and W. Hong, "Energy-efficient data organization and query processing in sensor networks," in *Proceedings of the 21st International Conference on Data Engineering (ICDE '05)*, pp. 157–158, New York, NY, USA, April 2005.

[4] X. Li, Y. J. Kim, R. Govindan, and W. Hong, "Multi-dimensional range queries in sensor networks," in *Proceedings of the 1st International Conference on Embedded Networked Sensor Systems (SenSys '03)*, pp. 63–75, Los Angels, Calif, USA, November 2003.

[5] Y. C. Chung, I. F. Su, and C. Lee, "Supporting multi-dimensional range query for sensor networs," in *Proceedings of the 27th IEEE International Conference on Distributed Computing Systems (ICDCS '07)*, pp. 35–44, Toronto, ON, Canda, June 2007.

[6] Y. Ren, V. Oleshchuk, and F. Y. Li, "A scheme for secure and reliable distributed data storage in unattended WSNs," in *Proceedings of the 53rd IEEE Global Communications Conference (GLOBECOM '10)*, Miami, Fla, USA, December 2010.

[7] W. H. Liao and C. H. Huang, "An efficient data storage scheme for top-k query in wireless sensor networks," in *Proceedings of the IEEE Network Operations and Management Symposium (NOMS '12)*, pp. 554–557, Maui, HI, USA, April 2012.

[8] Z. Ning, S. Shi, J. Li et al., "Multi-resolution data storage method for region query in sensor network," *Journal of Computer Research and Development*, vol. 49, no. 3, pp. 589–597, 2012.

[9] Z. Li, N. Xiao, and F. Chu, "A bouncing-track based data storage and discovery scheme in large scale sensor networks," *Journal of Computer Research and Development*, vol. 47, no. 11, pp. 1911–1918., 2010.

[10] Y. Ren, V. A. Oleshchuk, and F. Y. Li, "Optimized secure and reliable distributed data storage scheme and performance evaluation in unattended WSNs," *Computer Communications*, 2012.

[11] Z. Tao, Z. Gao, and Z. Lu, "Two new push-pull balanced data dissemination algorithms for large-scale wireless sensor networks," *Journal of Computer Research and Development*, vol. 45, no. 7, pp. 1115–1125, 2008.

[12] J. L. Bentley, "Multi-dimensional binary search trees used for associative searching," *Communications of the ACM*, vol. 18, no. 9, pp. 509–517, 1975.

[13] L. Xie, L. Chen, D. Chen, and L. Xie, "A decentralized storage scheme for multi-dimensional range queries over sensor networks," in *Proceedings of the 15th International Conference on Parallel and Distributed Systems (ICPADS '09)*, pp. 166–173, Shenzhen, China, December 2009.

[14] B. Karp and H. T. Kung, "GPSR: greedy perimeter stateless routing for wireless networks," in *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking (MOBICOM '00)*, pp. 243–254, August 2000.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 963508, 7 pages
http://dx.doi.org/10.1155/2013/963508

Research Article

Sensor Protocol for Roaming Bluetooth Multiagent Systems

Neungsoo Park,[1] Bijoy Kumar Mandal,[2] and Young-Ho Park[3]

[1] *Department of Computer Science and Engineering, Konkuk University, Seoul 143-701, Republic of Korea*
[2] *Computer Science & Engineering Department, NFET, NSHM Knowledge Campus—Durgapur, Durgapur 713212, India*
[3] *Department of Multimedia Science, Sookmyung Women's University, Seoul 140-742, Republic of Korea*

Correspondence should be addressed to Young-Ho Park; yhpark@sm.ac.kr

Received 29 January 2013; Accepted 4 April 2013

Academic Editor: Sabah Mohammed

Bluetooth is a low-cost, short-range wireless technology capable of providing many communication functionalities. However, Bluetooth does not support any sensor protocol which is related to a roaming and in which handoff occurs dynamically when a Bluetooth device is moving away from coverage of the network. If a device is losing its connection to the master device, there is no provision which transfers it to another master. Handoff is not possible in a piconet, as in order to stay within the network, a slave would have to keep the same master. Thus, by definition intrahandoff is not possible within a piconet. This research mainly focuses on Bluetooth roaming sensor technology and designs a sensor protocol which works in a roaming for Bluetooth multiagent system technology. The advantage of designing a roaming protocol is to ensure the Bluetooth enabled roaming devices can freely move inside the network coverage without losing its connection or break of service, in case of changing the base stations.

1. Introduction

Bluetooth is a low-cost, short-range wireless technology capable of providing many communication functionalities, ranging from wire replacement to simple personal area networking [1]. The range of Bluetooth network can be 0 to 100 meters. It is used to transmit both synchronous and asynchronous data. The bandwidth of the Bluetooth network, at physical layer, is 2.1 Mbit/s. Compared with other systems in the same frequency band, the Bluetooth radio hops faster and uses shorter packets. With respect to other wireless communication devices, Bluetooth connection can support both data and voice communications.

Bluetooth has two types of networks, piconets and scatternets. Piconet (PAN) consists of a master and one to seven active slaves. In a piconet, the devices share the same frequency-hopping spread spectrum (FHHS) channel, which is a transmission technology used in a local area wireless network. A scatternet is a collection of piconets and a connection node which links two piconets. This node can be simultaneously a slave of multiple piconets or a master in only one piconet. On the other hand, a Bluetooth device can be master in one and slave in other piconets, or slave in all piconets it is connected to [2].

An agent refers to an entity that functions continuously and autonomously in an environment in which other processes take place and other agents exist [3]. It can be a physical or software entity. Multiagent systems (MAS) involve a team of intelligent agents working together to accomplish a given task. The accomplishment of a task depends on the coordination of actions and behaviour of the agents. To achieve coordination between agents, communication can be used. Usually, multiagents operate in a close proximity. Thus, Bluetooth technology can be employed to set up the communication among agents.

This research mainly focuses on Bluetooth roaming sensor technology and designs a sensor protocol which works in roaming for Bluetooth multisystems technology. The advantage of designing a roaming protocol is to ensure that the Bluetooth enabled roaming devices can freely move inside the network coverage without losing its connection or break of service in case of changing the base stations.

The remainder of this paper is organized as follows. Section 2 explains Bluetooth protocol stack. Section 3

describes our methodology. Section 4 discusses how to connect Bluetooth devices. Section 6 illustrates data transfer operation. Section 7 explains how to connect slave which goes out of the piconet. Section 8 demonstrates handover old access point to new access points. Section 9 highlights conclusions and future work.

2. Bluetooth Protocol Stack

The Bluetooth protocol allows the authentication and encryption of a link at the same time. The security is controlled by the lower layers of the Bluetooth protocol. First, security level is assured by the fact that each Bluetooth module has a unique MAC (48-bits) address. In this way, when a scanner connects itself to a terminal with which it is twinned, the bar code is transmitted to this equipment. This process is shown in Figure 1.

The baseband layer is responsible for controlling and sending data packets over the radio link. It provides transmission channels for both data and voice. The baseband layer maintains Synchronous Connection-Oriented (SCO) links for voice and Asynchronous Connectionless (ACL) links for data. The SCO link is a symmetric point-to-point link between master and single slave in the piconet. SCO packets are not retransmitted, but ACL packets are, to ensure data integrity [5]. The Link Manager Protocol (LMP) uses the links set up by the baseband to establish connections and manage piconets. Responsibilities of the LMP also include authentication and security services and monitoring of service quality [5].

The Host Controller Interface (HCI) is the dividing line between software and hardware. The Logical Link Control and Adaptation Protocol (L2CAP) and layers above are currently implemented in software, and the LMP and lower layers are in hardware. The HCI is the driver interface for the physical bus that connects these two components. The L2CAP can be accessed directly by the application, or through certain support protocols provided to ease the burden on application programmers. Quality of Service (QoS) parameters are exchanged at this layer as well.

3. Methodology

The design process of a sensor protocol for roaming Bluetooth, equipped with MAS, will be carried out in two steps. The first step is to define roaming, when it should occur, for example, when an object will realize that it is going out of the range. And the second step is how to design a roaming sensor protocol. The main focus of this paper is to design a sensor protocol for roaming Bluetooth equipped with MAS.

3.1. Connection and Data Transfer on Roaming Devices. Here we assume that master is fixed unit connected to a fixed network, while slaves are mobile units. If a mobile unit loses the connection to the fixed unit it currently communicates with, a new unit is chosen to relay all the packets to the communication partner. For connection and data transfer between master and roaming slave, we need a sensor slave between them. For example, slave is a common unit between both piconets in Figure 2.

3.2. Maintain a Connection If Slave Is out of the Reach of Master. To know when slave will lose connection, RSSI (Received Signal Strength Indication) will be checked if it is in acceptable level. This means the roaming object is within the range of the network coverage. If the RSSI value of the moving object is less than the acceptable level, then it means the object is going out of range of the network coverage. That is the point where the object will start roaming and find the value of RSSI. There is a command used in Host Controller Interface (HCL) layer of the Bluetooth protocol stack get_link_quality [2]. This command enables us to measure the quality of RSSI on the basis of which the roaming occurs.

4. Mathematic Modelling

Consider an object which wants to move in x- and y-coordinating system using Bluetooth technology for communication. The position of the object with reference to x, y is $A(x, y)$ as shown in Figure 3.

Let r be the distance object from its signal source. Thus, after resolving r into the rectangular components, the horizontal and vertical component becomes

$$x = r\cos\theta, \tag{1}$$

$$y = r\sin\theta. \tag{2}$$

Now, if we square and add (1) and (2), we get the following formula:

$$\therefore r = \sqrt{(x^2 + y^2)}. \tag{3}$$

The position of the object with respect to x- and y-coordinating system is as follows:

$$\tan\theta = \frac{y}{x} \longrightarrow \theta = \tan^{-1}\frac{y}{x}. \tag{4}$$

Velocity can be defined as

$$\text{velocity} = \frac{\text{displacement}}{\text{time}}. \tag{5}$$

Thus, (2), (3), and (4) for X-Axis become (with respect to time)

$$v_x = \dot{x} = \frac{d(r\cos\theta)}{dt}, \quad \text{or, } v_x = \dot{x} = -r\sin\theta. \tag{6}$$

Similarly, for Y-Axis,

$$v_y = \dot{y} = r\cos\theta. \tag{7}$$

Differentiating (4) with respect to time,

$$\theta = \frac{\dot{x}\dot{y} - y\dot{x}}{x^2 + y^2}, \quad \text{angular velocity} = \frac{\text{angular position}}{\text{time}}. \tag{8}$$

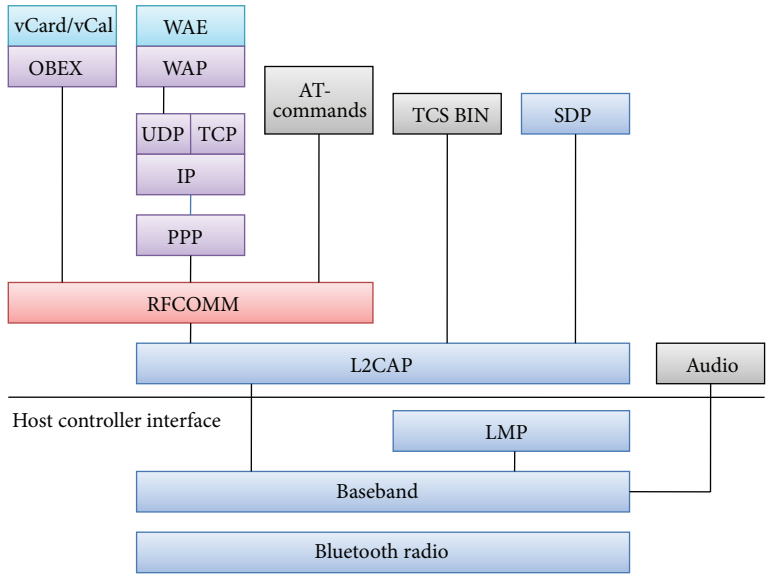


Figure 1: Bluetooth protocol stack model [4].

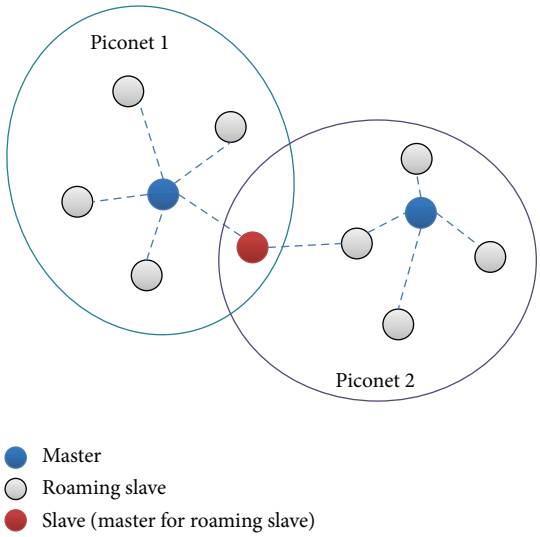


Figure 2: Data transfer from master to roaming slave.

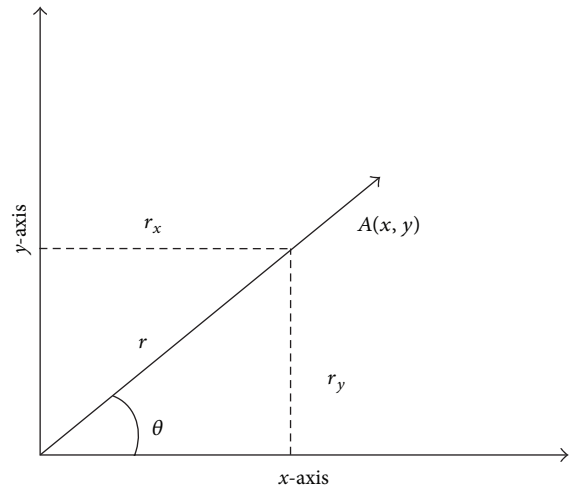


Figure 3: Graphical representation of object in x-y- coordinating system.

Equations (3) and (7) give us trajectory of the object in a piconet:

$$\therefore r = \sqrt{(x^2 + y^2)},$$
$$w = \theta \doteq \frac{x\dot{y} - y\dot{x}}{x^2 + y^2}. \tag{9}$$

5. Connecting Bluetooth Devices

A connection between two devices occurs in the following way: if any of the devices has no information about another/remote device, the inquiry, listing, and page procedure have to be carried out.

5.1. Inquiry Procedure. The inquiry procedure enables a device to discover devices in range and determine the addresses and clocks for the those devices. This procedure occurs in the following steps.

(1) Inquiry procedure involves a source device sending out inquiry packets and then receiving the inquiry reply. When a source device wishes to discover new devices, it enters the inquiry state, where it broadcasts inquiry packets (ID packets), containing the IAC, to all devices in range. It will send these using the inquiry hopping sequence. The device in the inquiry state can also receive inquiry replies (FHS packets); however, it will not acknowledge these packets.

(2) The destination device that receives the inquiry packets will have to be in the inquiry scan state to receive the inquiry packets.

(3) The destination device will then enter the inquiry response state and send an inquiry reply (own address and clock) to source device.

(4) After receiving inquiry reply, source device will create a record.

(5) Master will show a list of devices to a user. Table 1 shows an example of this list.

Suppose device 1 is a roaming 2 of slave 1 and device 2 is a roaming 1 of slave 2. Then master will focus on those devices which are needed for the connection, having minimum path from master to that device and all the devices.

(6) Master will send a list of devices, as in Step (5), to all the devices which are in the list after or during page procedure.

5.2. Listing Procedure. The listing procedure is carried out in followings steps as

(1) master will put its address and a few bits as a record in its database;

(2) during the inquiry procedure, slave will enter the inquiry response state and send an inquiry reply (its own address) to its master. After receiving address of each slave, master will add a new record to its database;

(3) now, slave 1 will create a record like in Step (1). During inquiry procedure of slave, roaming slave will enter the inquiry response state and send an inquiry reply (its own address) to slave. After receiving address of each roaming slave, slave will add a new record to its database as described in Step (2).

(4) all slaves will send their own record to the master as shown in Figure 4;

(5) master adds records to its own database.

5.3. Paging Procedure. With the paging procedure, actual connection can be established. The page procedure follows the inquiry procedure and listing procedure. Only the Bluetooth device address is required to set up a connection. Knowledge about the clock (clock estimate) will accelerate the set-up procedure. A unit that establishes a connection will carry out a page procedure and will automatically be a master of the connection. The procedure occurs as follows.

(1) A source device pages the destination device. Source device searches for another destination devices. The source device sends out a page packet (ID packet), using the page hopping sequence, to notify other devices that it wants to obtain destination device and/or their services.

(2) The destination device receives the page. A destination device listens for page trains containing its own device access code (DAC). When a destination device wishes to receive page packets, it enters the page scan mode. The scanning will follow the page hopping sequence. If a destination device receives a page packet, it will enter in slave response state.

(3) The destination device sends a reply to the source. Once a destination device has received its own DAC from the source (in the ID packet), it will enter this state. It will send a response message (its DAC again) to the source using the page response hopping sequence.

(4) The source device sends FHS packet to the destination device. When the source device has received a reply to its original page message, it will enter this state. It will then send FHS packet to the destination device using the page hopping sequence.

(5) The destination device sends its second reply to the source device. Once the destination device has received the FHS packet from the source, (Page Master Response State: Steps (3) and (4), the destination device will send a reply to the source (an ID packet containing the destination DAC).

(6) The destination device and source device then switch to the source channel parameters. When the source device has received the second reply (Page Slave Response State: Steps (3) and (5)), it knows that the destination device has received the FHS packet (Page Master Response State: Steps (3) and (4)). The source device is now the master of the destination device (the slave) and the destination device will switch to the source device's channel parameters. The destination device is now the slave of the source device (the master).

Table 1: User devices.

From User's list (selected these)	Background (in mobile)
Device 1	Device 1 addr + extra bits
Device 1	Device 2 addr + extra bits

The connection state starts with a poll packet sent by the master to verify that slave has switched to the master's timing and channel frequency hopping. The slave can respond with any type of packet. The master will send the page packets to slave. Using extra bits, slave detects the roaming slave and sends the page packets which it has got from master. When roaming slave wants to send any information to master, it will send this to slave. Then slave will forward the package to master.

6. Data Transfer: Packets

Each packet consists of 3 entities, the access code (68/72 bits), the header (54 bits), and the payload (0–2745 bits) as shown in Figure 5.

(1) Access code: access code is used for time synchronization, offset compensation, paging, and inquiry.

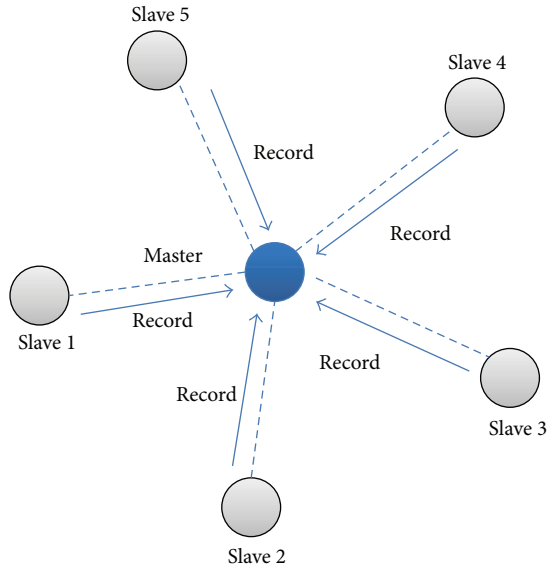


Figure 4: Slave sending its own record to master.

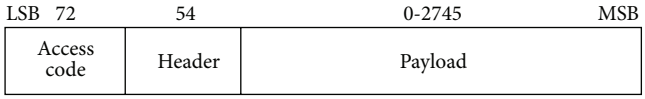


Figure 5: Packet Format.

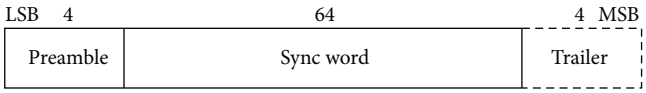


Figure 6: Access code format.

There are three different types of access code: Channel Access Code (CAC), Device Access Code (DAC), and Inquiry Access Code (IAC). The channel access code identifies a unique piconet, while the DAC is used for paging and its responses. IAC is used for inquiry purpose [6] as shown in Figure 6.

(2) Header: the header contains information for packet acknowledgement, packet numbering for out-of-order packet reordering, flow control, slave address, and error check for header [4] as shown in Figure 7.

(3) Payload: the packet payload can contain either voice field or data field or both. It has a data field; the payload will also contain a payload header [3].

7. Connecting Slave Which Goes out of the Piconet

To maintain security and connectivity in Bluetooth, another procedure will be added after paging procedure. Following paging procedure, connection will be created. Now the first package that will be sending before any data transfer will contain the list of devices going to get the same file. Reconnection is required only if a roaming slave 1, as shown in Figure 8, went out of piconet, and if it acts as a roaming slave in some other piconet.

If roaming slave 1 went out of the reach of its master (slave 1), then it will go for inquiry procedure. After inquiry procedure, it will get a list of devices with which it can create a connection. Now it will compare this list with the list it has gotten after this procedure and then find the common devices, that is, sensor in the list. If roaming slave 1 will get any device, then it will send a request to that device and go for paging procedure. Then, roaming slave 1 will update database and send a message to other devices to update their own records for the changes that have taken place in the scatternet. Otherwise it will show a warning message and disconnect itself from the scatternet. There might be multiple Bluetooth devices offering the same service, and most often the best one is the one from which we receive the strongest signal (because we will get a lower Bit Error Ratio and we will decrease the frequency of handovers). So, it will now find the device close to it. The flow chart of connection of slave is shown in Figure 9.

8. Handover Old Access Point to New Access Point

A Bluetooth device is connected with another Bluetooth device and suddenly moves out of range. After some time, the Bluetooth stack in the device declares the link to be dead and closes the LMP connection. At this point, the L2CAP used by the network traffic receives an event (LP_DisconnectInd) and closes the connection down. The higher layer is kept in suspended mode until the roaming procedure is completed. This procedure is explained in [7].

At this point, the Bluetooth device performs an inquiry. If the inquiry does not find any access point, the node continues with inquiry, up to the point where it times out and closes

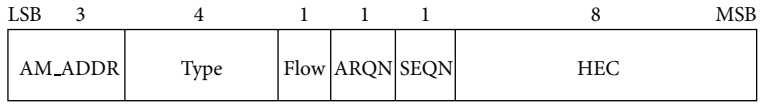


Figure 7: Header Format.

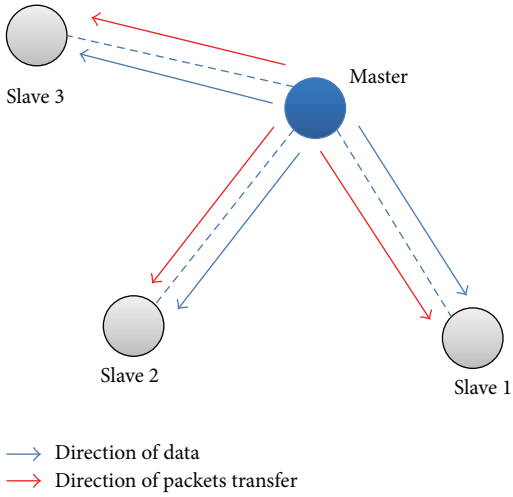


Figure 8: Data transfer in devices.

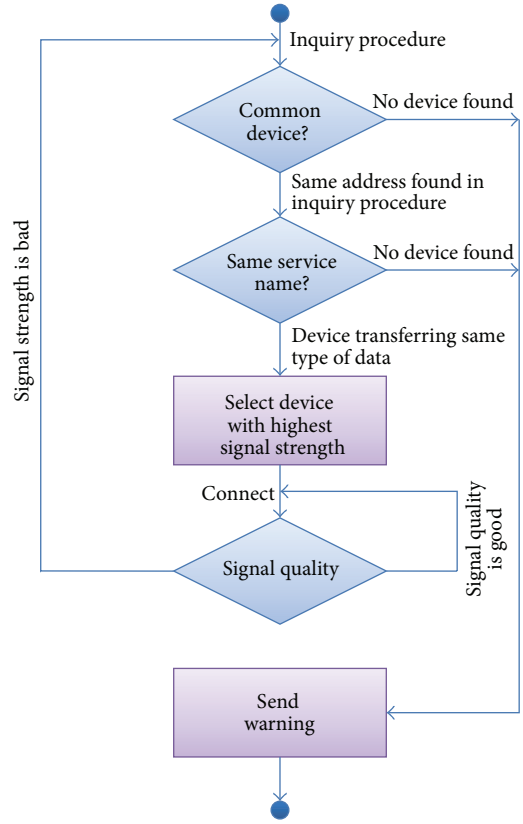


Figure 9: The flow chart of connection of slave.

down the higher layer (either PPP or BNEP). If the Bluetooth device finds another Bluetooth devices, then if the "Service Name" of this Access Point is the same as the one of the previous access point, the device connects to it. If the "Service Name" is different and if the device has discovered other access points in the inquiry process, the device should try to connect to those access points. If the device cannot find an access point with the same "Service Name," it can either try to connect to one of the other access points, continue doing inquiry, or return a failure message to the user.

While the node is performing its handover, the old access point will still continue to receive and buffer packets for the roaming node, but cannot deliver them. When the node is associated with the new access point, the new access point starts to receive packets for that node at this point, but does not grab earlier packets. In other words, all the packets sent from the infrastructure while the node was performing the handover are not received by the node. Those packets will need to be resent from the source, which will consume time. On the other hand, the old access point has stored most of those packets, and they will just wait in its buffer. When the old access point receives the deregistration message from the new access point, it can simply resend all the packets in its buffer associated with this node. The new access point will pick them up and deliver them to the node. The access point will collect the BR address from the slave and go for an inquiry procedure to find the old access point of this slave. Using extra bits, it will be easier to find the location of old access point.

9. Data Sharing in Slaves

Data sharing is one of the most important methods to increase the speed during data transfer and reduce the load of master. In this case, two or more slaves will share the data between them, so that source will not have to send the data individually to all the slaves. The flow chart of data sharing method in slaves is shown in Figure 10.

10. Conclusion

It is a widely held belief that a lack of the ability to create a "proper" Bluetooth pan-piconet network is one major drawback of the present incarnation of the Bluetooth specification. However, it must be noted that Bluetooth was originally designed as a low-cost/low-power system. The designing of a roaming protocol is to ensure the Bluetooth enabled roaming devices can freely move inside the network coverage without losing its connection or break of service in case of changing the base stations. The discovery process to establish connection in Bluetooth network is time-consuming process.

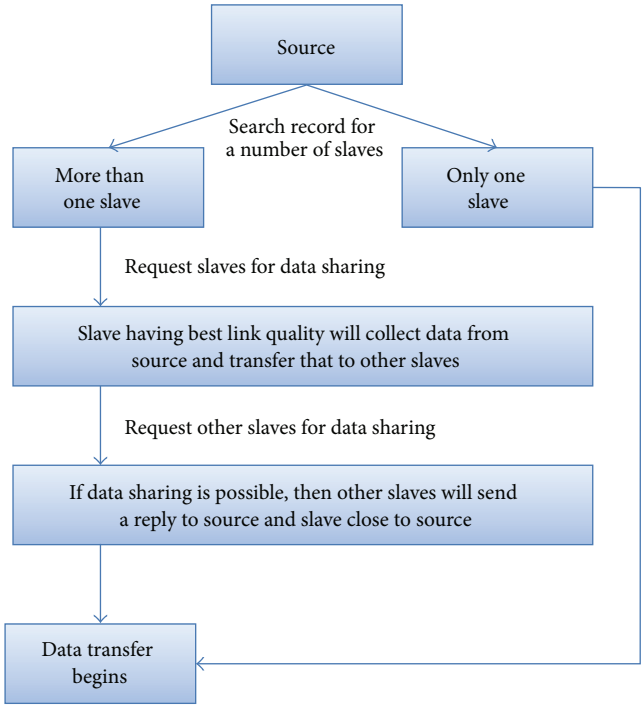


Figure 10: Flow chart of data sharing in slaves.

Using this roaming protocol, we also can reduce the time. Since Bluetooth is a short-range wireless communication protocol, we shall design protocol for a long-range wireless communication.

Acknowledgment

This work was supported by the IT R&D Program of MKE/KEIT (10041854, Development of a smart home service platform with real-time danger prediction and prevention for safety residential environments).

References

[1] F. Mazzenga, D. Cassioli, A. Detti, I. Habib, P. Loreti, and F. Vatalaro, "Performance evaluation in bluetooth dense piconet areas," *IEEE Transactions on Wireless Communications*, vol. 3, no. 6, pp. 2362–2373, 2004.

[2] H. Pals, Z. R. Dai, J. Grabowski, and H. Neukirchen, "UML-based modeling of roaming with bluetooth devices," in *Proceedings of the 1st Hangzhou-Lubeck Conference on Software Engineering (HL-SE' 03)*, pp. 1–7, 2003.

[3] G. Mejia and L. J. Mahecha, "A multi-agent system approach for the supply chain of fresh products in Bogota, Columbia," Joint Pricing and Inventory Decisions for Perishable Products.

[4] N. Minar and M. Tarique, "Bluetooth security threats and solutions: a survey," *International Journal of Distributed and Parallel Systesm*, vol. 3, no. 1, pp. 127–148, 2012.

[5] Y. T. Chung, *Bluetooth announcement system [Dissertation]*, Universiti Teknologi Malaysia.

[6] Bluetooth SIG, "Specification of the Bluetooth System, Core, Part B," Version 1. 1, 2001, http://www.bluetooth.com.

[7] J. Tourrilhes, "BlueTooth roaming proposals," Working document submited to the BlueTooth PAN working group, 2000, http://www.hpl.hp.com/personal/Jean_Tourrilhes/Papers/.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 973621, 15 pages
http://dx.doi.org/10.1155/2013/973621

Research Article

Wireless Virtual Multiple Antenna Networks for Critical Process Control: Protocol Design and Experiments

Stefano Savazzi

IEIIT Institute of the Italian National Research Council (CNR), Milan, Italy

Correspondence should be addressed to Stefano Savazzi; stefano.savazzi@ieiit.cnr.it

Received 26 November 2012; Revised 11 March 2013; Accepted 11 March 2013

Academic Editor: Carlos Ramos

Wireless telemetry systems for remote monitoring and control of industrial processes are now becoming a relevant topic in the field of networked control. Wireless closed-loop control systems have stricter delay and link reliability requirements compared to conventional sensor networks for open-loop monitoring and call for the development of advanced network architectures. By following the guidelines introduced by recent standardization, this paper focuses on the most recent technological advances to enable wireless networked control for tight closed-loop applications with cycle times below 100 ms. The cooperative network paradigm is indicated as the key technology to enable cable replacing even in critical control applications. A cooperative communication system enables wireless devices placed at geographically separated locations to act as a virtual ensemble of antennas that creates a virtual multiple-antenna-distributed system. A proprietary link-layer protocol/based on the IEEE 802.15.4 physical layer has been developed and tested in an indoor environment characterized by non-line-of-sight (NLOS) propagation and dense obstacles. The measurements obtained from the testbed evaluate experimentally the benefits (and the limitations) of cable replacing in critical process control.

1. Introduction

The increasing demand of oil and gas supplies frequently requires the design of very large production and processing plants over remote locations with harsh environmental conditions and challenging logistics. The adoption of cabling to fully interconnect machines and monitor/control large number of processes is becoming unfeasible due to the high fluctuations of installed industrial wiring costs [1].

In networked control systems, the controller and the plant are connected via a digital communication channel of limited bandwidth [2]. A cable-based networked control architecture is considered in Figure 1: the sensors monitor any plant activity and periodically forward the digital measurements (y_k) to a remote controller. Based on these observations, the remote controller computes a sequence of control messages (u_k) according to a given policy and sends them to the actuators over the feedback channel. Upon retrieval of the controller messages, the actuators apply appropriate control signals to adjust the plant state. For several process control applications like semiconductor manufacturing, tooling machines, production of nanomaterials, and so forth, the determinism of data transfer is a key issue, and the cycle time (i.e., round trip time) is a critical parameter to guarantee process stability [2].

The adoption of wireless technology in critical industrial applications is still rather limited: it is generally acknowledged that to allow a wider adoption of wireless networks in an industrial context, some substantial technology innovation is required either based on new physical layer solutions or on different approaches at the upper protocol layers [3]. Industrial networks typically require low-jitter-sampling period for monitoring, high integrity data delivery of critical messages, automatic reconfiguration, and usage of redundancy in case of communication failures. The most representative application cases where the wireless technology is adopted can be found in [4–6]. The available commercial wireless systems for industrial control and monitoring predominantly use the ISM bands at 2.4 GHz and prevent the adoption of wireless in emergency actions and tight process control loops. Today, commercial battery-operated systems are based on the IEEE 802.15.4 standard and enable data to

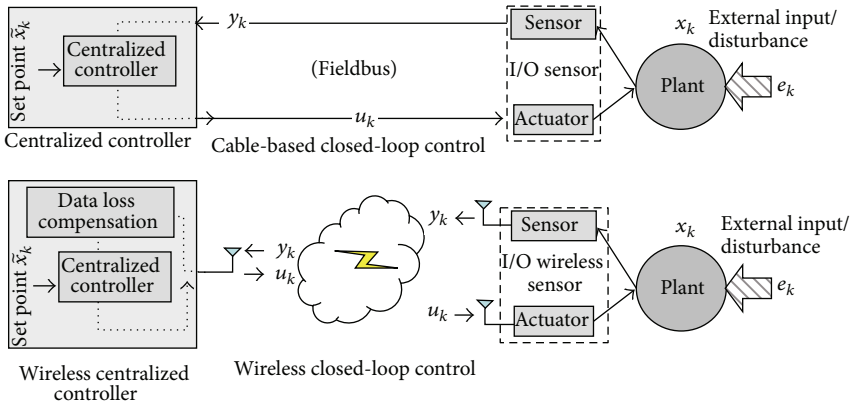


Figure 1: Cable-based versus wireless closed-loop control systems.

be transmitted at a typical rate of 250 kbit/s (and scaling up to 2 Mbit/s by disabling direct-sequence spread spectrum functions), with up to a maximum of 10 dBm output RF power to meet the RF regulations for hazardous environments. The IEEE 802.15.4 physical layer also constitutes the basis for the wireless HART [7] and ISA100.11a [8] industry standard protocols.

This paper focuses on the most promising technologies to support the next generation wireless control systems designed for tight closed-loop applications. The wireless communication system used to transmit observations and control messages must guarantee a minimum quality of service in order for the system to be controllable. Some recent works in this domain have highlighted the relation between the unreliability of the transmission channel and the controllability of the system stability, showing that a strict relation exists between the transmission channel characteristics and the unstable poles of the open-loop system [9–11]. In these works, the impact of the noisy transmission channel is mostly considered for the feedback loop with the assumption of simple additive white Gaussian (AWGN) channel model. Without looking at more realistic scenarios where fading is the main impairment of the communication, those approaches are, thus, very prone to failure in realistic contexts.

In this paper, we evaluate experimentally the impact of fading channels on the controllability of the closed-loop wireless system. In particular, it is envisaged here that the incorporation of the cooperative network paradigm [12] into future wireless system standardization will allow cable replacing in tight closed-loop control applications with cycle time below 100 ms [2]. Cooperative communication systems emulate the transmission and the reception of data on a (virtual) antenna array, thus, creating a virtual and distributed multiple antenna array network [13]. To highlight the potential of such systems, a proprietary link-layer protocol tailored for closed-loop process control applications has been developed on top of an existing IEEE 802.15.4 compliant PHY/MAC layer radio stack. Real-time process control has been tested in an indoor environment with non-line-of-sight (NLOS) propagation and dense obstacles. Results from the testbed measurements confirm that cooperative communication is a promising enabling technology for the next generation critical wireless control systems as it provides clear performance advantages compared to classical network architectures in terms of link reliability and closed-loop stability performance. Analysis of experimental data reveals that the configuration and planning of the wireless control system should account for the stability properties of the plant process. This imposes a substantial redefinition of conventional wireless network deployment and design methods.

2. Wireless Closed-Loop Control Networking

In what follows, we consider a control network with clock-driven sensing. The focus of the analysis is thus on a scenario where the wireless network is constrained to periodically monitor and control the process state being subject to unpredictable disturbance. The output sensor in active state is periodically sampling a continuous signal $\mathbf{y}(t) \in \mathbb{R}^m$ with period T_s (reporting rate) to obtain the time vector series $y_k = \mathbf{y}(t_k)$, $t_k = kT_s$. Discrete signals $y_k \in \mathbb{R}^m$ provide an observation of the plant state vector $x_k \in \mathbb{R}^q$. The plant model for process observations is described in discrete-time state-space form:

$$\begin{aligned} y_k &= \mathbf{C} \times x_k + n_k, \\ x_k &= \mathbf{A} \times x_{k-1} + \mathbf{B} \times u_k + \mathbf{D} \times e_k + w_k, \\ u_k &= \mathbf{G}\left(x_{k-1}, x_{k-2}, \ldots \mid \widetilde{x}_k\right), \quad \forall k. \end{aligned} \tag{1}$$

At time t_k, the plant vector state x_k is a function of the previous states x_{k-1}, the feedback control variable $u_k \in \mathbb{R}^v$, and the external random input process $e_k \in \mathbb{R}^v$ acting as an external nonstationary disturbance. The feedback control u_k is generated by the controller on every new received process observation. Control message follows a generic control law function $\mathbf{G}(\cdot)$ that depends on all the previous vector states $x_{k-1}, x_{k-2}, \ldots$ estimated from the corresponding noisy observations $y_{k-1}, y_{k-2}, \ldots$. The purpose of the controller is to stabilize the system by balancing the external input disturbance and minimizing the deviation of plant states x_k from the desired stable set points indicated here by $\widetilde{x}_k$. Given that the focus is on wireless control performance assessment, it is assumed that any observation y_k reliably transmitted over the wireless link provides a full state measurement

of x_k and is affected by a scaling factor modeled as a full-rank matrix $\mathbf{C}$. The instrument AWG noise $n_k \in \mathbb{R}^m$ with $n_k \sim \mathcal{N}(\mathbf{0}, \sigma_n^2\mathbf{I})$ includes quantization and other unwanted effects. The noise term $w_k \in \mathbb{R}^q$ accounts for the state disturbance and is modeled as independent AWG noise with $w_k \sim \mathcal{N}(\boldsymbol{\mu}_w, \sigma_w^2\mathbf{I})$ so that $\mathbf{x}_0 = \boldsymbol{\mu}_w$.

The round-trip latency T_{RT} is a critical parameter for process control, and it is defined as the time between the sampling (and transmission) of the observation y_k and the successful decoding of the feedback control message u_{k+1} by the remote actuator. A networked control system that satisfies the stabilizable properties needs two additional conditions to guarantee closed-loop stability: (i) the observation y_k and the feedback control u_k must be successfully decoded by the respective parties; (ii) the tolerable round-trip latency is such that $T_{\mathrm{RT}} \leq T_s$.

In this paper, the main focus is on cable replacing for control systems requiring $T_{\mathrm{RT}} < 100$ ms. This is a reasonable choice to address most industrial control applications [2]. The case for highly critical control (e.g., motion control) that requires cycle times $T_{\mathrm{RT}} < 10$ ms is not considered here as still too challenging for implementation over current low-power wireless technology.

2.1. System Model. The development of robust network designs requires accurate modeling of radio propagation to account for the random fluctuations of the received signals due to fading impairments [14]. To simplify the reasoning, we assume the output sensor and the actuator to be colocated and referred to as input/output sensor (I/O sensor). Extension to a more general model is straightforward. Both the controller and the I/O sensor are deployed at fixed locations over the plant and equipped with a radio device characterized by a single omnidirectional antenna transceiver and a limited battery energy supply mainly used for the transmission, reception, and processing of data. Transmission of measurements s_k (over uplink) and feedback control u_k (over downlink) is subject to half-duplex constraint so that it occurs in different time slots and satisfies the round trip delay constraint T_{RT}. Let $d_{\mathrm{I,C}}$ be the distance between the I/O sensor I and the controller C; the probability of successful closed-loop control P_c is modeled by outage probability

$$P_c = \Pr\left[\min\left\{\gamma_{\mathrm{I,C}}, \gamma_{\mathrm{C,I}}\right\} \geq \beta\right], \tag{2}$$

where $\gamma_{\mathrm{I,C}}$ is the Received Signal Strength (RSS) measured by the controller C over the uplink, while $\gamma_{\mathrm{C,I}}$ is the RSS observed by the actuator I over downlink. β models the sensitivity of the receiver and depends critically on hardware implementation and on modulation of signals. The RSS $\gamma_{a,b}$ for a wireless link (a, b) depends on deterministic components (transmitter/receiver distance, height from ground, and obstruction size/position) and on random components due to multipath-fading impairments [15]. An effective statistical description of fading channel terms can be obtained by Weibull distribution [16].

Assuming statistical independence between uplink and downlink RSS fluctuations, successful control probability P_c can be rewritten as the product of the success probabilities over uplink and downlink

$$P_c = \underbrace{\Pr\left[\gamma_{\mathrm{I,C}} \geq \beta\right]}_{\text{Uplink: Sensor} \to \text{Controller}} \times \underbrace{\Pr\left[\gamma_{\mathrm{C,I}} \geq \beta\right]}_{\text{Downlink: Controller} \to \text{Actuator}}. \tag{3}$$

This assumption is also confirmed by measurements over the 2.4 GHz spectrum (see Section 5).

2.2. Closed-Loop Control Performance Metrics. The probability of successful control P_c (or equivalently the packet loss rate) is a good indicator of networked control quality as stability is primarily ruled by packet loss rate [17]. Given that it is important to develop an understanding of how much loss the control system can tolerate before observing instability [18], an additional measure to characterize the stability properties of the process is the stability interval $T_{\mathrm{stability}}$. The stability interval measures how infrequent feedback information is needed to guarantee that the system remains stable, even if subject to packet drops [19]. A large packet loss rate (or small enough P_c) causes frequent interruptions of closed-loop control while if the duration of those interruptions exceeds $T_{\mathrm{stability}}$, that is, as observed during deep fades, then the process states might experience large deviations from the desired stable set points or become unstable for even longer communication interruptions.

A convenient metric used to evaluate process stability is

$$P_{\mathrm{stability}} = \Pr\left[\left\|x_k - \widetilde{x}_k\right\| \leq \delta\right] \tag{4}$$

that measures the probability that the deviation of process states x_k from the stable set-points $\widetilde{x}_k$, that is, caused by random packet losses, lies below an accuracy parameter $\delta > 0$. This factor δ indicates a critical condition for HW instrumentation that might cause costly losses for the plant operator. Stability probability $P_{\mathrm{stability}}$ is computed over K consecutive loops where the process can be reasonably assumed as ergodic.

3. Cooperative Communication for Critical Networked Control

The emerging area of cooperative communications suggests that it is worthwhile to explore the potential of advanced network architectures where the classic constraints valid for wired communications are relaxed [12]. The *cooperative link* abstraction consists of separate radios encoding and transmitting messages in coordination. Both information-theoretic (see, e.g., [20, 21]) and experimental analyses [22, 23] showed that under specific conditions on the propagation environment, a cooperative system could achieve similar performance to colocated multiantenna systems. A cooperative network architecture has the potential to be less sensible to isolated wireless link failures, compared to non-cooperative architectures, as it creates a virtual distributed antenna network consisting of multiple paths where the same information is spread to maximize path redundancy (spatial or cooperative diversity).

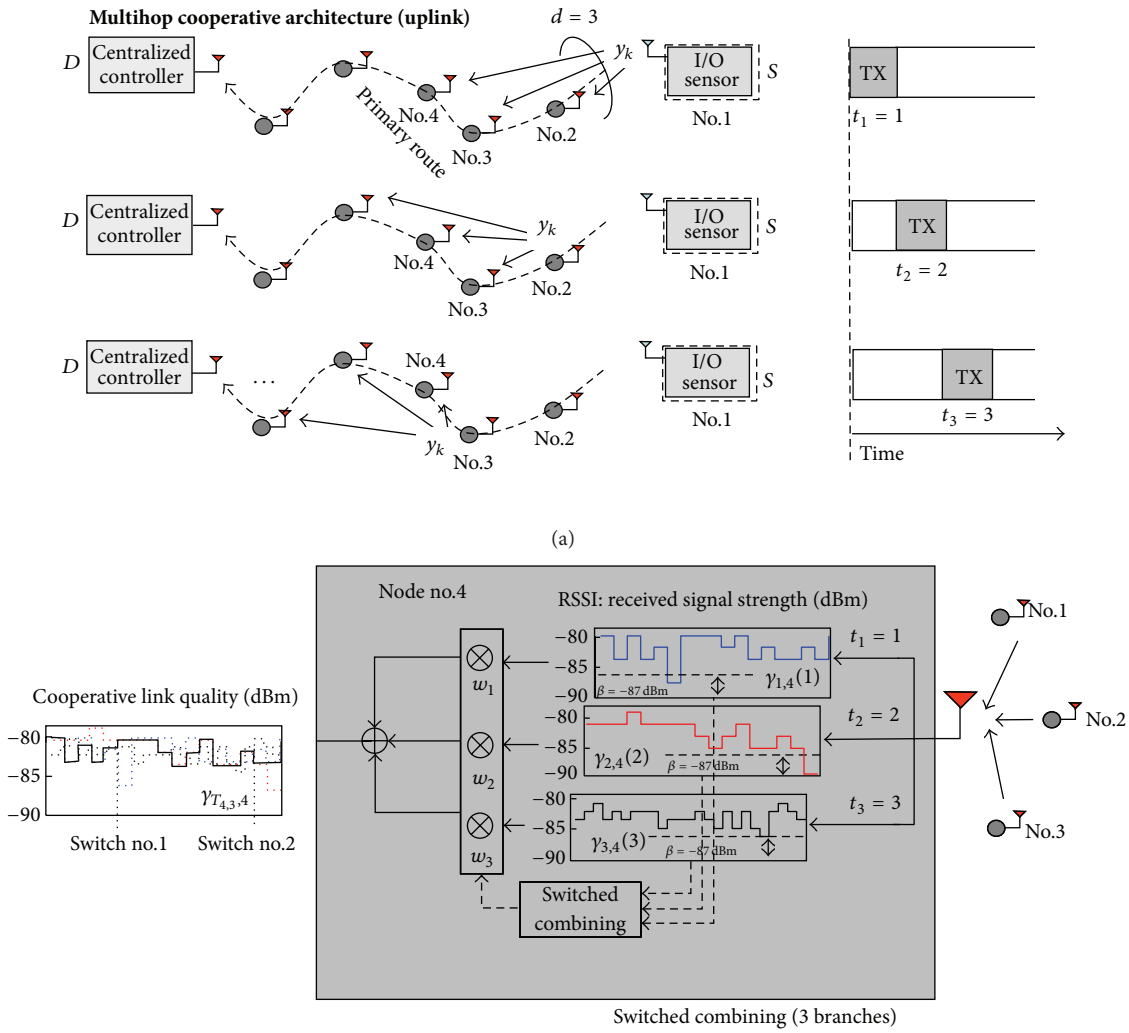


FIGURE 2: Data propagation (uplink) over the primary route by multihop cooperative architecture with diversity $d = 3$ (a). Switched combining example at node no.4 (b) uses $d = 3$ replicas of observation y_k while signal strengths are taken from measurements at 2.4 GHz. In switch no.1 link $(1, 4)$ is replaced by link $(2, 4)$ while after switch no.2 link $(3, 4)$ is chosen.

3.1. Multihop Cooperative Link Modeling. To introduce the problem of multihop-cooperative link performance modeling, let the wireless control network in Figure 2 be represented by a set of randomly distributed nodes within a specific area. A sequence of messages is continuously transmitted by a *source* node S to a *destination* node D over an optimal "connection oriented" unicast route path $\mathcal{R}$ (primary route) that involves M intermediate nodes relaying data to destination D. Ordering of nodes is labelled as $\mathcal{R} = \{S, 1, 2, \ldots, M, D\}$ where the source node S and the destination node D take the role of I/O sensor and centralized controller for uplink, while their roles are reversed over downlink.

The propagation of the messages is based on time division access and it is illustrated in Figure 2(a). The *multihop-cooperative* architecture improves the reliability of multihop message passing along the primary route by implementing a chain of consecutive cooperative transmissions [13]. At time slot $t = 1$ (for convenience the time slots are numbered as for the nodes), the process observation is originated from I/O sensor source S and relayed at time $t = 2$ from node 2 and so on. Similarly, after the destination is reached, the same message propagation is initiated now by the centralized controller acting as the source node for backward propagation of the feedback control message. In general, for each transmitting node $k \in \mathcal{R} \setminus \{D\}$, there are up to d subsequent nodes in the route that are overhearing. Therefore, the kth receiver has up to d copies of the same message during d subsequent time slots that experience statistically independent fluctuations of the received signal strength and can be incrementally combined to exploit the *cooperative diversity* order of d. The cooperative set of nodes $\mathcal{T}_{k,d}$ that are transmitting towards the terminal k as part of the cooperative link $(\mathcal{T}_{k,d}, k)$ are defined as $\mathcal{T}_{k,d} = \mathcal{T}^U_{k,d}$ with $\mathcal{T}^U_{k,d} = \{k-d, \ldots, k-1\} \subset \mathcal{R}$ for uplink and as $\mathcal{T}_{k,d} = \mathcal{T}^D_{k,d}$ with $\mathcal{T}^D_{k,d} = \{k+d, \ldots, k+1\} \subset \mathcal{R}$

for downlink. For practical system design, a useful bound to the probability of successful control is

$$P_c = P_c(d) \approx \underbrace{\prod_{k\in\mathcal{R}\setminus\{S\equiv I\}} \Pr\left[\gamma_{\mathcal{T}^U_{k,d},k} \geq \beta\right]}_{\text{Uplink: } \Pr\left[\gamma_{I,C}\geq\beta\right]} \times \underbrace{\prod_{k\in\mathcal{R}\setminus\{S\equiv C\}} \Pr\left[\gamma_{\mathcal{T}^D_{k,d},k} \geq \beta\right]}_{\text{Downlink: } \Pr\left[\gamma_{C,I}\geq\beta\right]}, \tag{5}$$

being the product of successful probabilities over all the cooperative links $(\mathcal{T}^U_{k,d}, k)$, $(\mathcal{T}^D_{k,d}, k)$ with $k \in \mathcal{R} \setminus \{S\}$ defined for uplink and downlink, respectively. The approximation holds for large enough Signal-to-Noise Ratio (SNR) [21]. Unlike conventional multihop message passing, each kth receiver combines the RSSs measured over the d links involved in collaborative transmission. The term $\gamma_{\mathcal{T}_{k,d},k}$ measures the quality of the virtual cooperative link $(\mathcal{T}_{k,d}, k)$ and depends on the selected combining scheme as illustrated in the following section.

3.2. Selection and Switched Combining. The selection combining technique can be employed to exploit the redundancy made available by the cooperative network architecture. The selection combining scheme allows each receiver to decode only the message copy originated from the link that experienced the highest instantaneous RSS. From (5) the d combining weights w_h with $h \in \mathcal{T}_{k,d}$ over the d links are such that

$$\gamma_{\mathcal{T}_{k,d},k} = \sum_{h\in\mathcal{T}_{k,d}} w_h \gamma_{h,k} = \max_{h\in\mathcal{T}_{k,d}} \gamma_{h,k}, \tag{6}$$

where $w_h = 1$ if and only if $h = \arg\max_{h\in\mathcal{T}_{k,d}} \gamma_{h,k}$ and zero otherwise. The probability of successful control (5) can be bounded as

$$P_c = P_c(d) > \prod_{k\in\mathcal{R}\setminus\{S\}} \Pr\left[\max_{h\in\mathcal{T}^U_{k,d}} \gamma_{h,k} \geq \beta\right] \times \Pr\left[\max_{h\in\mathcal{T}^D_{k,d}} \gamma_{h,k} \geq \beta\right], \tag{7}$$

where $\Pr[\max_{h\in\mathcal{T}_{k,d}} \gamma_{h,k} \geq \beta] = 1 - \prod_{h\in\mathcal{T}_{k,d}} \Pr[\gamma_{h,k} < \beta]$.

An alternative option to selection combining is the switched combining scheme that allows the device to switch to the best link only if the previously chosen connection (e.g., with node h) undergoes a deep fade such that $\gamma_{h,k} < \beta$. The implementation of the switched combining scheme is illustrated in the example of Figure 2 (at (b)) for cooperative diversity order $d = 3$. Although selection combining outperforms switched combining in terms of average performance (same outage probability performance is observed at high SNR), switched combining requires only a single RF chain to serve all the cooperative links and is practical enough for implementation over low-power devices.

4. Virtual Multiple Antenna MAC Protocol Design

Most existing works on cooperative communication focus on various aspects at physical layer while the advantages of the proposed schemes are often demonstrated by using an information theoretic approach. Many results are therefore based on asymptotically large data frame length assumption and usually ignore the upper layer overhead required to set up, synchronize, and coordinate a cooperative system [24]. MAC protocol design for cooperative communications has recently been a hot research topic [25]. Cooperative MAC protocols can be classified into proactive and reactive schemes [24]: proactive schemes always provide one (or more) prearranged and optimal partner(s) serving as relay node for the source node [26]; reactive schemes prescribe that cooperative transmission is initiated only when a negative acknowledgement (NACK) message is received (see, e.g., [27]). Extensive work has been reported in the literature relating to MAC design based on modifications of the distributed coordination function (DCF) of IEEE 802.11 standard. Several protocol designs have been proposed for single and multiple relay networks employing both fixed relaying assignments [26–28] and dynamic [27–29] assignments (relay selection). In these papers, both decode and forward (DF), amplify and forward (AF) and coded cooperation strategies have been investigated. Some attempts in the literature have been made towards the definition of MAC specifics to enable cooperative communication over IEEE 802.15.4 networks (see, e.g., [30]), although the topic is still considered an open issue.

The proposed cooperative MAC protocol depicted in Figure 3 is defined on top of the IEEE 802.15.4-2011 PHY layer and it is based on a proactive scheme. The network architecture consists of three components, detailed as follows.

(i) The *Centralized Controller* manages a low-power radio interface for two-way communication with the remote I/O sensor and acts as a translator over the wired network. The centralized controller transmits set-points $\tilde{x}_k$ and computes control commands u_k to guarantee the global stability of the plant.

(ii) The *virtual controllers* are the additional infrastructure used to emulate the virtual antenna array system. The virtual controllers take the dual role of cooperatively receiving the plant observations from the I/O sensor and replacing the centralized controller when its direct link with the actuator experiences any degradation. The virtual controllers act as leaf nodes for propagating the decisions made by the central controller and have no permission to generate new set points. In case of consecutive packet drops, they can replace the centralized controller to secure local stability and data loss compensation.

(iii) The *I/O sensor* is the low-power input/output field instrument that interacts with the plant behavior generating process observations y_k and applying control commands u_k.

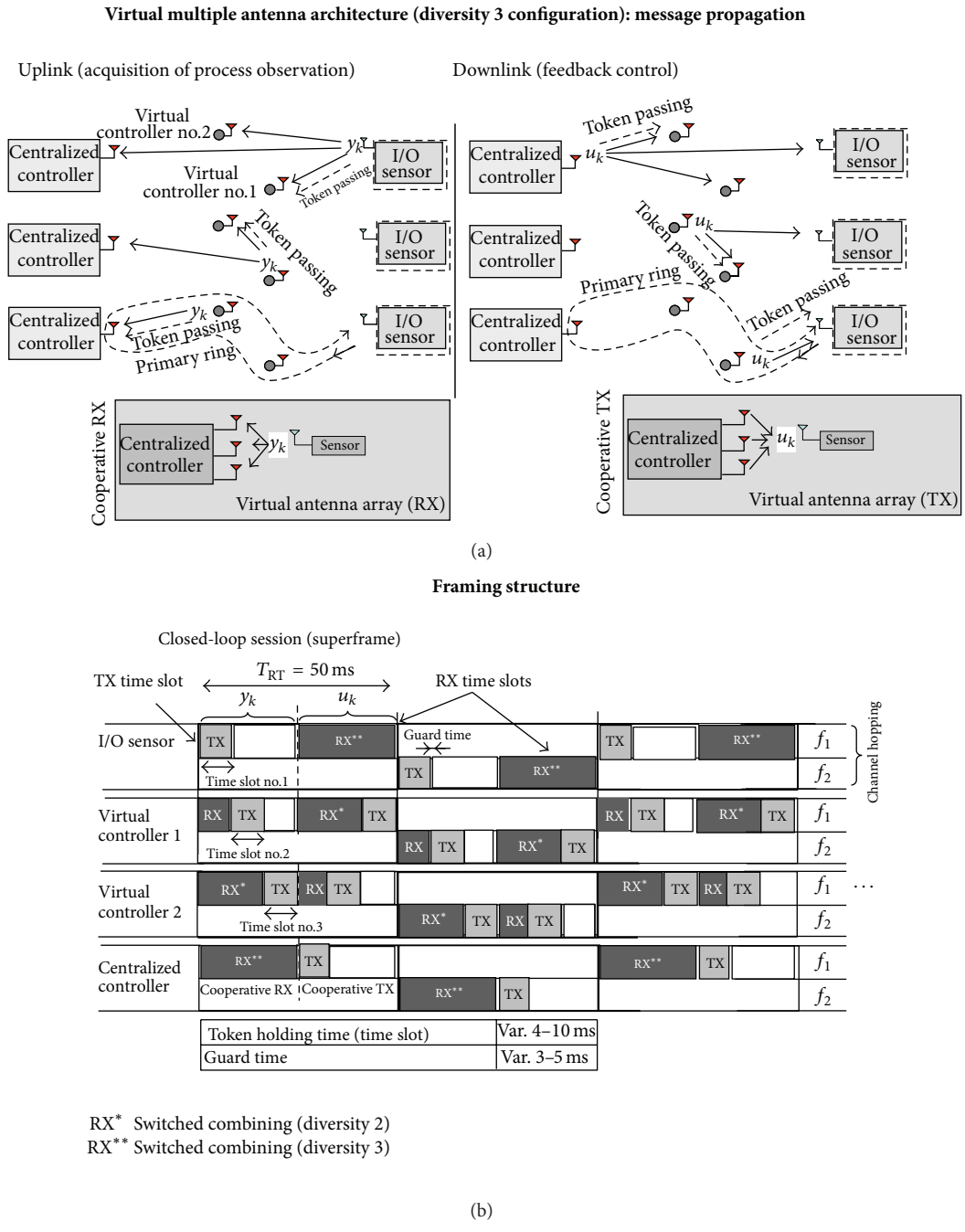


FIGURE 3: Virtual multiple antenna array system architecture: message passing over the multihop cooperative transmission chain (a) for uplink (left side) and downlink (right side): the token message passing is also superimposed (dashed arrows). Virtual antenna arrays are shown at (a) and provide (in this example) a cooperative diversity of $d = 3$. Framing structure for timed-token MAC (b): the example refers to the case of $M = 2$ virtual controllers while token holding times, guard times are illustrated in the table at (b). Channels used for FH are $f_1 = 2.425$ GHz and $f_2 = 2.455$ GHz, respectively.

The message-passing scheme and the framing structure depicted in Figure 3 refer to a system deploying $M = 2$ virtual controllers with maximum allowed cycle time T_{RT}. An analogous framing structure can be defined for an arbitrary number M of virtual controllers. A time division duplex system is employed to separate uplink and downlink. Transmissions are organized into consecutive superframes consisting of $2(M + 1)$ time-slots of length T separated by guard times of length ΔT to compensate for residual clock misalignments. Each superframe contains one closed-loop session (or cycle time) of T_{RT} sec. A closed-loop session starts with the transmission of the available measurement y_k and stops when the feedback control u_k is received and applied to the plant. The transmission of the noisy process

sample y_k is delayed by the I/O sensor until the assigned time slot is obtained. The measurement is then propagated by the M virtual controllers towards the centralized controller according to the multihop cooperative network architecture described in Section 3. When the measurement is received by the centralized controller, the new sample is used as input to generate the new control message u_k. Similarly, as for the process samples, the control message is then propagated over the downlink using the assigned time slot.

The proposed cooperative network protocol is based on a frequency-hopped, timed-token message passing scheme (see Section 4.1). Devices implement an ad hoc cooperative link control policy to handle the switched combining of the signal replicas and network synchronization (see Section 4.2). Data loss compensation is applied to address the residual impairments observed over the cooperative wireless channel (see Section 4.3).

4.1. Medium Access Control Sublayer. The medium access control sublayer implements a channel frequency hopping (FH) over consecutive superframes. FH is commonly adopted in industrial networks as it allows the system to be less susceptible to interference, providing some additional protection against eavesdroppers. Within each superframe, the medium access control uses a timed-token message passing protocol on top of the multihop cooperative network architecture described in Section 3. The timed-token protocol has been also proposed to enforce real-time on wired/wireless Profibus and industrial Ethernet networks, overriding the native collision-based multiple access [31]. A token message is multiplexed with information data to form a frame (token frame) and visits all the devices on every closed-loop session to synchronize the cooperative transmissions. The token holding time is bounded to the duration T of one time slot to satisfy the round trip delay requirement T_{RT}.

During MAC configuration, the network is organized into a logical *primary ring* connecting the I/O sensor to the centralized controller and vice versa. The primary ring is a two-way routing path connecting the controller with I/O sensor through the M virtual controllers. The configuration of the primary routing path is optimized as it is based on radio planning. In complex environments like refinery or power plants, the use of the 3D model during the design phase is also crucial to maximize radio-planning accuracy in order to limit any rework to a percentage which is in line with a regular installation of a wired system [15].

The amount of cooperative diversity d is decided based on the behavior of the process in open loop (further details are given in Section 6): the selected cooperative diversity limits the number virtual MIMO links that can be combined. When the cooperative diversity order is chosen, the centralized controller assigns to devices one time slot (TX time slot) for transmission and up to d time slots (RX time slots) for receiving redundancy over the virtual multiple-antenna links (in uplink and downlink).

4.2. Cooperative Link Control and Synchronization. The cooperative architecture imposes a redefinition of conventional logical link control designs. An additional level of abstraction compared to multihop networks should be defined to efficiently manage and control the "cooperative link" as the set of physical links involved in collaborative transmission. The proposed cooperative link control implements a switched combining scheme configured to estimate the RSS during the assigned RX time slots and switch to the best link if the measured RSS goes below the threshold β. The purpose of switched combining is to enforce the real-time constraint by avoiding the use of error control methods based on explicit acknowledgements [4].

The periodic token-passing procedure plays also a crucial role to guarantee device synchronization [32]. When a device overhears a new token message, it computes the misalignment between the expected and the actual time of arrival of the token packet. This information is then used to predict the next time-to-token visit time $T_{\text{token-visit}}$ (and thus the beginning of the assigned time slot). Every new timing update for $T_{\text{token-visit}}$ must account for the particular path over which the token frame is received: given that the token frame is received by device k and transmitted by device $h \in \mathcal{T}_{k,d}$, the next time to token visit is computed as

$$T_{\text{token-visit}} = \Delta T + (h - k - 1) \times (T + \Delta T) + \tau_{k,h}, \quad (8)$$

where $\tau_{k,h}$ is the random misalignment (in number of OQPSK symbols, with duration 16 μs) measured by node k between the expected and the actual time of arrival of the token frame from node h.

4.3. Data Loss Compensation. Any residual data loss over the cooperative links might result in missing plant measurements at the centralized controller. The virtual controllers and the centralized controller are thus designed to predict the missing sample y_k based on the p previous samples $\mathbf{y}_{k-1} = [y_{k-1}, \ldots, y_{k-p-1}]^T$ according to the linear predictor

$$\widehat{y}_{k|k-1} = \mathbf{a}_p^T \mathbf{y}_{k-1}. \quad (9)$$

For a stationary process, the minimum mean square error (MMSE) predictor is obtained by letting

$$\mathbf{a}_p = \mathbf{C}_p^{-1} \mathbf{r}, \quad (10)$$

where $\mathbf{C}_p = E[\mathbf{y}_k \mathbf{y}_k^H]$ and $\mathbf{r} = E[y_k \mathbf{y}_{k-1}^H]$ are the covariance and cross-correlation of the stationary process observations, respectively.

Gradient-based linear prediction is a common choice in predictive model-based control [2] as model parameters for linear regression are estimated from data samples $\mathbf{y}_k$ without a priori information about the statistical behavior of the process. Prediction is obtained by

$$\mathbf{a}_p = \mathbf{N}_p \times \mathbf{z}, \quad (11)$$

where $\mathbf{N}_p = \mathbf{P} \times (\mathbf{P}^T \mathbf{P})^{-1}$, $\mathbf{P} = [\mathbf{t}_k, \mathbf{1}]$, $\mathbf{t}_k = [(p-1)T_s, (p-2)T_s, \ldots, 0]^T$, and $\mathbf{z} = [pT_s, 1]^T$.

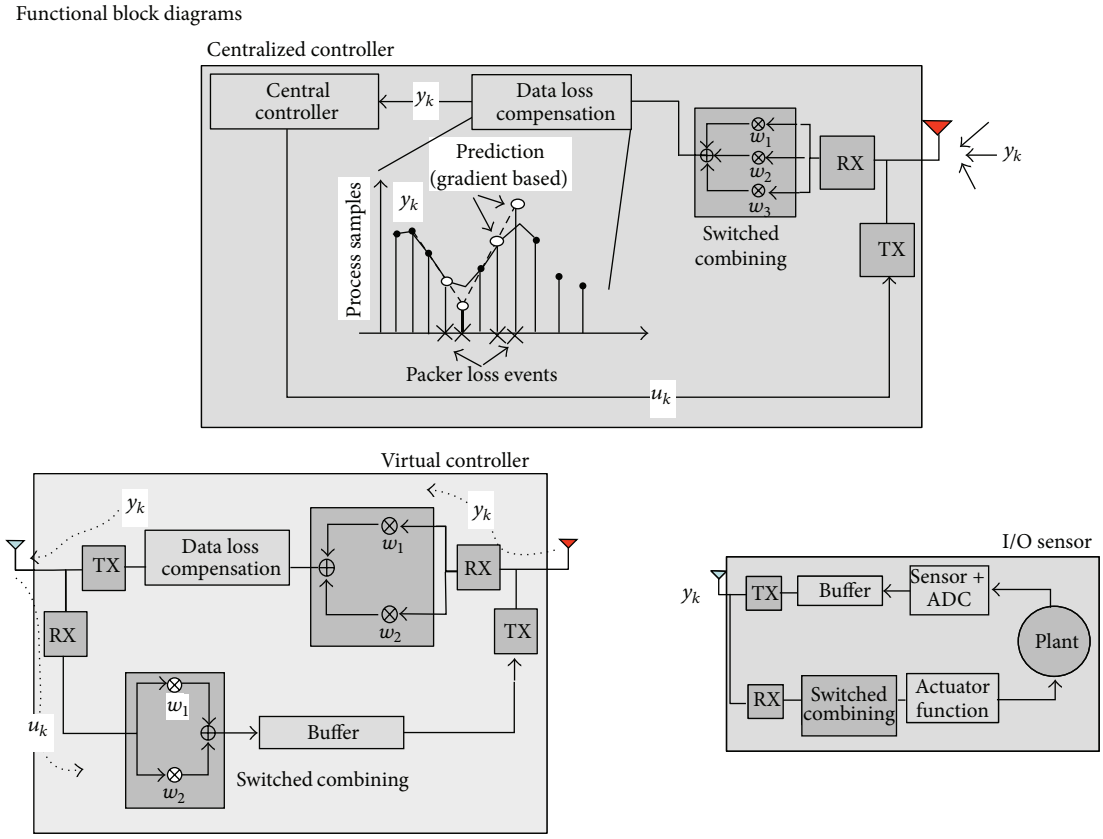


FIGURE 4: Functional block diagram for centralized controller, virtual controller, and I/O sensor.

5. Wireless Critical Process Control System Implementation: Case Study

In the proposed experimental setup, the virtual multiple antenna protocol specifics are implemented over battery-powered motes based on the low-power CC2420 single-chip 2.4 GHz IEEE 802.15.4 compliant [33] with radio transmit power set to $P_T = 1$ mW. The IEEE 802.15.4 PHY layer allows the use of 16 channels for FH where each channel occupies an effective bandwidth of 2 MHz with center frequency separation of 5 MHz.

The RSS Indicator (RSSI) is used to assess the link quality for switched combining with $\beta = -87$ dBm [16]. The RSSI provides an estimate of the signal power by energy detection over 8 consecutive offset quadrature phase shift keying (O-QPSK) symbols, corresponding to a duration of 128 μs. The RSSI is quantized using 8 bit/sample and stored in the CC2420 RSSI_VAL register.

As depicted in Figure 3, the duration of one closed-loop session equals the superframe length of $T_{\mathrm{RT}} = 50$ ms; a guard time of $\Delta T = 3$ ms among consecutive superframes is adopted. The frequency hopping is performed over consecutive closed-loop sessions: the hopping pattern periodically switches among the IEEE 802.15.4 channels with center frequencies 2.425 GHz and 2.455 GHz, corresponding to the channel numbers 15 and 21, respectively. Frequency-hopping requires on/off radio switching and introduces a latency of ~2 ms. The selected channels are marginally influenced by cross-tier interference that originated from WiFi or Bluetooth modules [34].

The superframe is divided into slots of fixed length of $T = 5$ ms. The IEEE 802.15.4 slotted CSMA-CA access implemented by the devices is modified so that the back-off function is disabled. An energy scan to detect cross-tier interference (by clear channel access CCA) is performed at the beginning of the assigned slot; in case the channel is sensed as free, the transmission of the token frame is performed with the acknowledgement option disabled. The token frame structure is based on the IEEE 802.15.4 beacon frame type and contains the control message u_k (for downlink) or the actual/predicted process sample y_k (for uplink). Additional information is embedded in each frame to identify (i) the closed-loop session; (ii) the current set-point $\tilde{x}_k$; (iii) the device type and position within the primary ring; (iv) the channel offset for frequency hopping; (v) the selected diversity order d.

5.1. System Implementation. In what follows, the application-specific system implementation is detailed by looking at each network component separately (see also Figure 4).

(i) Centralized Controller. The centralized controller is equipped with a low-power 8-bit AVR microcontroller implementing a linear state-feedback controller such that $u_k = \mathbf{K}\hat{x}_{k-1}$,

where $\widehat{x}_k = \mathbf{C}^{-1} y_k \, \forall k$ while feedback gain matrix $\mathbf{K}$ is designed to achieve the desired closed-loop pole locations (see, e.g., [35]). When a new measurement is received either from the virtual controllers or the I/O sensor, a notifying indication event is generated by the MAC layer to inform the controller that a new control message is required. Control message is then forwarded by the centralized controller over the assigned time slot. The centralized controller generates new set points $\widetilde{x}_k$ and acts as a translator over the wired network by communicating with a device serving as gateway node. Even if the chosen proportional control policy is fairly simple compared to conventional industrial process control systems [2], it is useful to highlight the potential benefits of the cooperative architecture.

(ii) Virtual Controller. On every new closed-loop session, the virtual controller is designed to receive and combine up to d copies of the signal encoding the sensor measurement over uplink and up to d copies of the signal carrying the control message over downlink. The RSSI is used as a metric to select the message copy to decode by switched combining (Section 4.2). In case of missing process observations, the gradient-based data loss compensation function (see Section 4.3) is implemented: similarly as for process observations, the predicted sample $\widehat{y}_{k|k-1}$ is now forwarded to the centralized controller over the assigned multihop cooperative links. In case of missing control messages, the virtual controller replaces the centralized controller to guarantee the stability of the set point (received before losing communication with the centralized controller). It therefore implements a linear state feedback control policy using the same feedback gain matrix $\mathbf{K}$ of the centralized controller.

(iii) I/O Sensor. The AVR microcontroller is used to emulate the transducer and the actuator functions of the field instrument by generating the simulated process observations obtained from the discrete-time state-space plant model in (1). The observations $y_k = x_k + n_k$ provide a noisy representation of the process states. The sampling time of the process is set to $T_s = 60$ ms: sampling process is implemented using a timer obtained by the system clock sourced by an external oscillator. Each observation is encoded before radio transmission using a 16 bit/sample. The I/O sensor uses a buffer of finite length that stores the available sample before transmission over the assigned time slot: the samples which do not belong to the current step are discarded. Any new control message received either from the virtual controllers or the centralized controller during a closed-loop session generates a notifying indication that activates the actuator functions. A plant state adjustment is thus simulated according to model (1): any state adjustment influences the upcoming process sample without introducing significant delay.

5.2. Experimental Activity. An example of a single hop (a) and of a virtual multiple-antenna-based (b) closed-loop control is depicted in Figure 5: the purpose is to assess process stability by visual inspection of plant variables with respect to accuracy threshold δ. In this example, the noisy observations y_k of one process state are visualized over a time window of

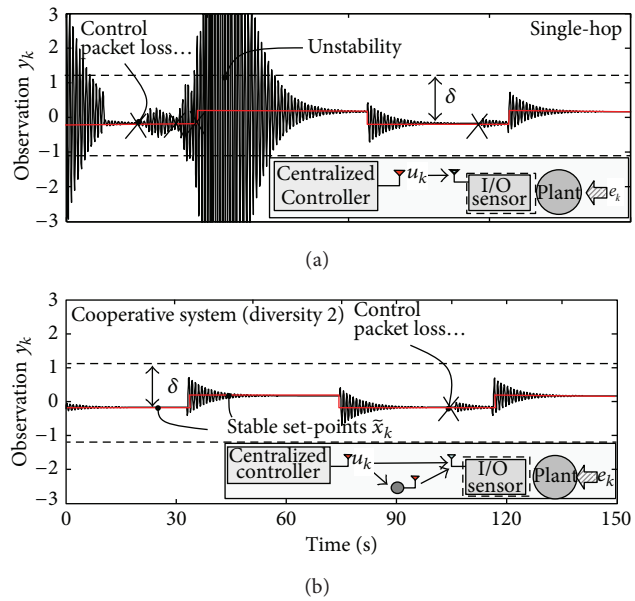


Figure 5: Example of wireless process control by single hop (a) and cooperative networking (b). Process observations are taken from simulated plant model (Model A, see Figure 6).

150 s. External input e_k (1) randomly switches among two set points on every 30 sec on average to emulate a nonstationary disturbance. The stable set-points $\widetilde{x}_k$ are depicted in solid red lines and depend on the external input disturbance. In this example, the use of a single-hop network architecture is not sufficient to guarantee stability while the virtual double-antenna option provides a clear advantage.

For the experiments, the considered indoor environment consisted of two rooms separated by a wall with 10 cm thickness. Up to 7 people were moving inside each room, and this causes random fluctuations of the radio signals. For all devices, the antenna height from ground is 1 m; the harsh radio environment was made of metallic objects (e.g., coaxial cabling, monitors/PCs, tubes for air conditioning, etc) responsible for additional attenuations. This is a worst case scenario as compared with typical industry standard installation designs that recommend 2 m height from the ground [36]. The centralized controller sends control messages to the I/O sensor placed in the adjacent room at a distance of 16 m (see topology superimposed on the floor plan in Figure 6). This specific setting is designed to assess the impact of NLOS propagation on the performance of closed-loop control.

For the proposed architecture, we considered the deployment of a single ($M = 1$ with diversity $d = 2$) and a pair ($M = 2$ with diversity $d = 3$) of virtual controllers. The performance of single-hop and multihop architecture are also evaluated. The *single-hop* scheme implements a standard ARQ policy where the retransmissions are subject to timing constraints and are thus confined within the time division-framing structure of Figure 4 with $T_{\mathrm{RT}} = 50$ ms. The multihop scheme requires the installation of a wireless repeater that implements decode and forward relaying [21].

Closed-loop control stability is evaluated over two state-space discrete time plant models: these are referred to as

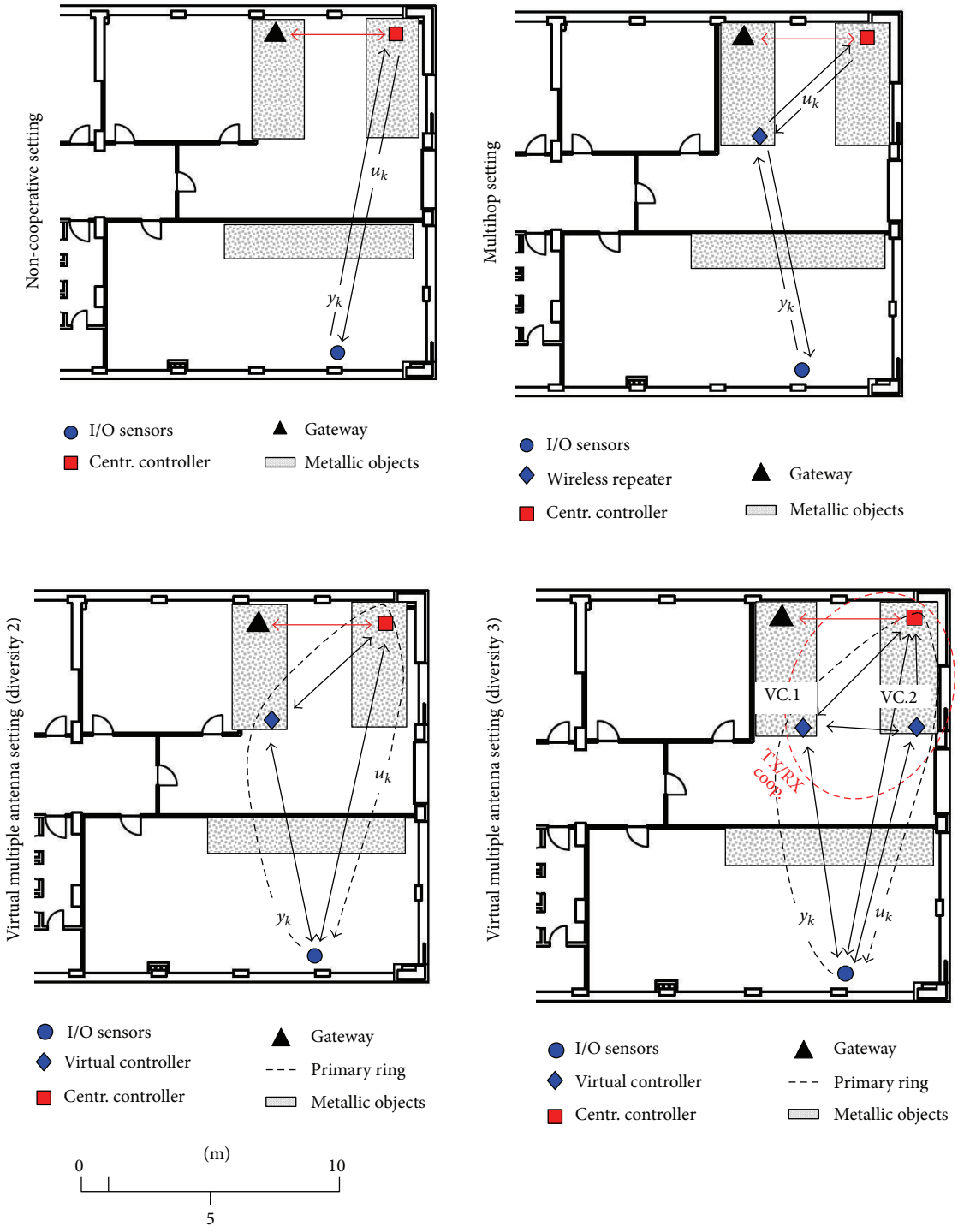


FIGURE 6: Floor plan map of the environment for experimental activity over 2.4 GHz. Network topologies for all settings are also superimposed.

model A and model B, respectively. Locations of the unstable open-loop and desired closed-loop poles are $0.85 \pm 0.625j$ and $0.85 \pm 0.5j$, respectively for Model A, while for model B these are $1.1 \pm 0.837j$ and $0.95 \pm 0.01j$. To analyze the impact of wireless propagation on closed-loop control performance, the stability interval $T_{\text{stability}}$ is evaluated (defined in Section 2.2). Stability interval defines the tolerable duration of the wireless link interruption (e.g., for N consecutive packet drops) above which the deviations from stable set-points $\tilde{x}_k$ are too large compared to accuracy δ in (4). Analysis over the first

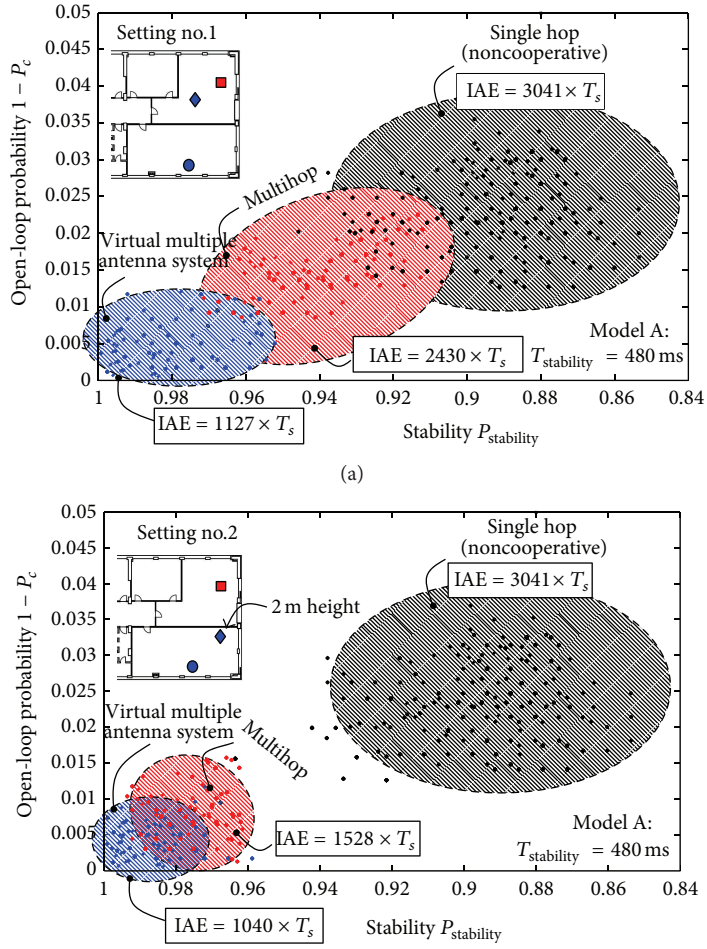


FIGURE 7: Closed-loop control performance for Model A with stability interval $T_{\text{stability}} = 480$ ms (corresponding to 8 consecutive packet losses). Each point maps to an open-loop probability $1 - P_c$ and stability $P_{\text{stability}}$ computed over 20 minutes of real-time control. Network topology "setting no.1" is depicted in (a). The case for optimal deployment of the virtual controller is shown in "setting no.2" at (b). Observed average IAE over $K = 20000$ consecutive control loops is also superimposed for each case.

configuration (model A) shows that up to $N = 8$ consecutive packet losses, corresponding to $T_{\text{stability}} = 480$ ms, are still tolerable in practice to stabilize the system dynamics. Instead, analysis over the more challenging plant model B shows that any link interruption above $T_{\text{stability}} = 180$ ms, corresponding to $N = 3$ consecutive packet losses, makes the system dynamics highly unstable.

Performance of closed-loop control is depicted in Figure 7 for plant model A and in Figure 8 for model B, respectively. For each setting, continuous real-time control is tested over a period of 5 days on average. In both figures, each point maps to the average open-loop probability $1 - P_c$ with P_c defined in (2) and the process stability $P_{\text{stability}}$ (4) observed over a time window of 20 minutes corresponding to $K = 20000$ process samples. Tolerable deviation from the stable set-points is based on feedback gain and chosen here as $\delta = \varsigma \times \max \|\tilde{x}_k\|$ with $\varsigma = 1/2$ so that the maximum deviation of measured state x_k from stable set-point $\tilde{x}_k$ lies below the 50% of the maximum range $\max_k \|\tilde{x}_k\|$.

In Figure 7, we compare the single-hop, the multi-hop, and the cooperative settings configured with a single virtual controller (with diversity $d = 2$). In Figure 7(a), the virtual controller is deployed in the same room of the centralized controller (setting no.1). This case is often typical in industrial settings where the I/O sensors are deployed in hazardous areas and require IP66/67 certification while the installation of additional infrastructure in the same area might not be allowed. In Figure 7(b), the virtual controller is now deployed in the same room of the I/O sensor (setting no.2) as this is the best choice for network planning to minimize the packet loss probability over the two-hop route. For both settings, the tolerable open-loop probability for 99% stability $P_{\text{stability}}$ should lie below 10^{-2} ($P_c > 0.99$). Only the cooperative architecture can guarantee such a high level of

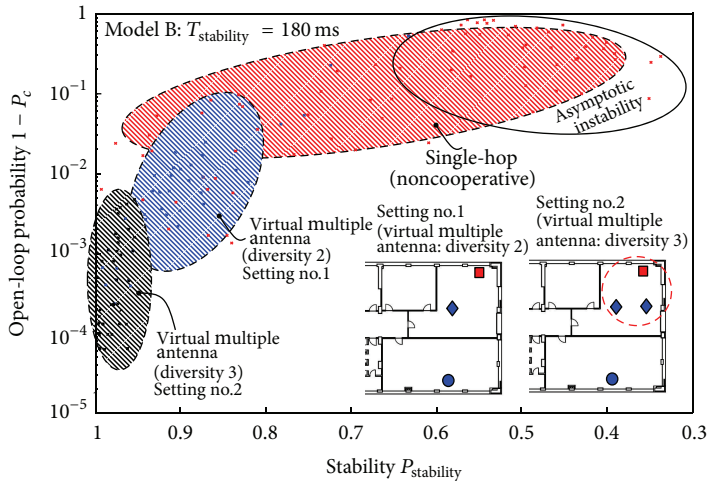


FIGURE 8: Performance analysis of closed-loop control for plant Model B with $T_{\text{stability}} = 180$ ms (corresponding to 3 consecutive packet losses).

reliability. The multihop architecture is highly sensible to the relay deployment as accurate network planning (if allowed) provides significant performance improvements as observed in Figure 7(b). For all the considered settings, the observed average integral absolute error [2] (IAE) over K consecutive control cycles IAE $= \sum_{k=0}^{K-1} \|x_k - \tilde{x}_k\| T_s$ is also superimposed and confirms the benefits of the proposed architecture.

The more challenging plant model B is analyzed in Figure 8; the performance of the single-hop scheme is compared with the cooperative system configured for $M = 1$ and $M = 2$ virtual controllers with cooperative diversity $d = 2$ and $d = 3$, respectively. The small stability interval $T_{\text{stability}}$ tolerated by the more critical plant model B suggests the use of 3 virtual antennas, thus configured for $M = 2$ virtual controllers, with cooperative diversity $d = 3$. This is the only viable solution to guarantee an average open loop probability below 10^{-3} ($P_c > 0.999$) for the desired 99% stability level.

6. Virtual Multiple Antenna System Design

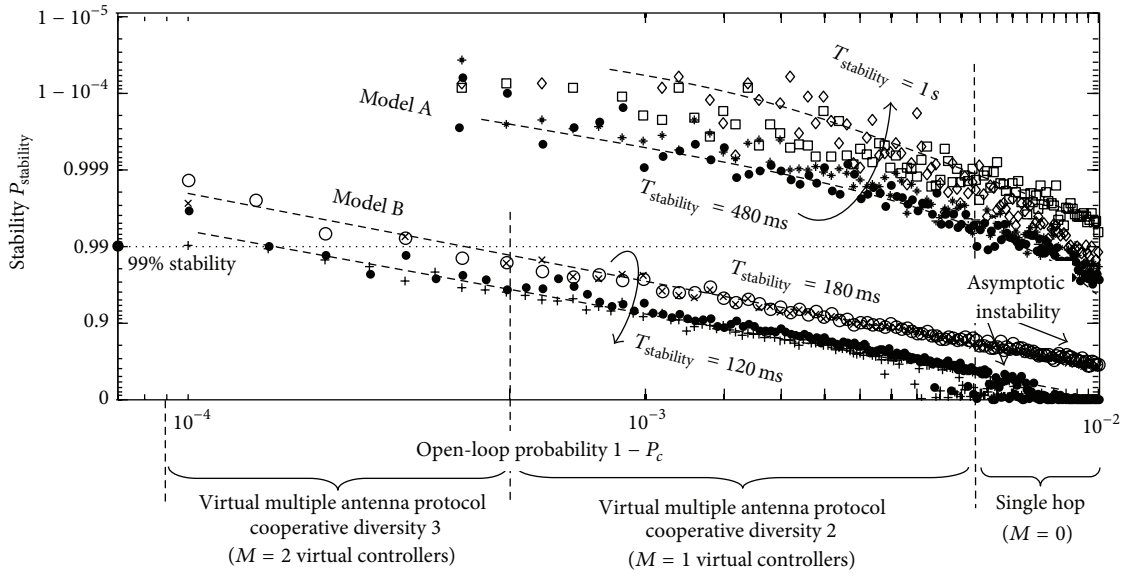
As described in the previous section, the stability interval $T_{\text{stability}}$ characterizes the behavior of the process in open loop. In addition, it defines a tolerable level of success probability P_c of closed-loop control above which the system can be considered as stable for all practical purposes: the lower the interval T_{stabilty}, the larger the tolerable success probability P_c. The choice of the cooperative diversity d and of the number of virtual controllers M for cooperative network planning should therefore account for these key design parameters. The purpose of this section is to highlight the factors that mostly influence the protocol configuration with special focus on the choice of the cooperative diversity d (Section 6.1) and its impact on the energy consumption (Section 6.2).

6.1. Cooperative Diversity Design. The proposed approach to the design of the cooperative diversity is to fix a required stability probability (here 99%) and to numerically choose the cooperative diversity to meet this stability constraint, thus limiting the number of consecutive packet drops accordingly. The required diversity therefore depends on the stability interval $T_{\text{stability}}$ of the considered plant model.

To allow for general insights, a simulation tool has been developed to assess the stability of the control system for varying plant models characterized by different values for the tolerable stability interval $T_{\text{stability}}$. In Figure 9, the control stability $P_{\text{stability}}$ is analyzed for varying open-loop probabilities $(1 - P_c)$ and plant models. Plant processes are indicated by different markers and experience different stability intervals $T_{\text{stability}}$ to model low ($T_{\text{stability}} = 1$ sec) to highly unstable ($T_{\text{stability}} = 120$ ms) behaviors. For each setting, the cooperative diversity d is chosen to guarantee the desired open-loop probability for 99% stability (dashed line). The required cooperative diversities d and open-loop probabilities are also reported in the table at the bottom as a function of $T_{\text{stability}}$.

The analysis clearly shows that the use of single and multihop architectures is reasonable for supervised control with $T_{\text{stability}} \geq 1$ sec where up to $N = 16$ consecutive packet drops are still tolerable to maintain stability. The cooperative scheme designed for diversity $d = 2$ is a reasonable option for process control with $T_{\text{stability}} = 480$ ms. Finally, the system configured for diversity $d = 3$ confirms as a promising option to support critical control with $T_{\text{stability}} \leq 180$ ms, for example, where no more than $N = 3$ consecutive packet drops are allowed.

6.2. Design Considerations for Battery-Powered Devices. In this section, the average power absorbed by the virtual controller device on every control cycle is computed. Notice that the virtual controllers experience the longest activity cycle as they employ selection combining over uplink and downlink. The purpose is to highlight relevant considerations for network lifetime prediction. To allow for general insights, the power consumption is modeled as a function of the required cooperative diversity d. The power absorption measurements



Virtual multiple antenna system design for 99% stability

Process stability time (max. number of consecutive packet drops)	Cooperative diversity	Desired open loop probability
$T_{\text{stability}}$ = 120 ms (2 consecutive packet drops)	$d = 3$	$<10^{-4}$
$T_{\text{stability}}$ = 180 ms (3 consecutive packet drops)	$d = 3$	$<2 \times 10^{-3}$
$T_{\text{stability}}$ = 480 ms (8 consecutive packet drops)	$d = 2$	$<5 \times 10^{-2}$
$T_{\text{stability}} \cong$ 1 s (16 consecutive packet drops)	Single hop	$<10^{-2}$

FIGURE 9: Stability $P_{\text{stability}}$ versus open-loop probability $1 - P_c$. Different markers refer to plant models with stability intervals $T_{\text{stability}}$ = 120 ms/1 sec. Models A and B are included as special cases. Performance of the proposed virtual multiple antenna system for different diversity configurations correspond to different open-loop probabilities as observed in the experiments (see Figures 7 and 8). (Bottom table) Cooperative diversity design for 99% stability: 4 plant models are considered with corresponding stability times $T_{\text{stability}}$, desired open-loop probability, and cooperative diversity.

are taken from the IEEE 802.15.4 compliant transceivers used during the experimental activity and specified for 2.7 V/3.3 V operation.

For a given slot duration T (token holding time) and closed-loop session T_{RT}, the virtual controller is designed to keep the radio transceiver active for receiving and combining up to $2d$ messages (sensor observations and control messages, resp.). Two additional slots are used for relaying messages over uplink and downlink. The average power consumption per control cycle P_{loop} is thus proportional to the selected diversity order d as

$$P_{\text{loop}} = d \times \frac{2P_{\text{rx}}(T + \Delta T)}{T_{\text{RT}}} + \frac{2P_{\text{tx}}T}{T_{\text{RT}}} + \left(1 - \frac{2T(d+1)}{T_{\text{RT}}}\right) P_{\text{sleep}}, \tag{12}$$

where P_{tx} = 62.7 mW is the average power absorbed (in milliwatts) during transmission at 3.3V while P_{rx} = 56.1 mW is the power absorbed in receiving mode. Power draw in sleep mode is 82.5 μW: during sleep mode the internal oscillator and RAM must be in active state (memory hold). The power absorbed by a virtual controller configured for diversity $d = 2$ is P_{loop} = 33.2 mW, while for diversity order $d = 3$ is 40% larger as P_{loop} = 45.9 mW. These results highlight the inherent trade-off between maximizing the control reliability (that requires a high spatial redundancy and a long duty cycle) and the network lifetime (that requires a long sleep cycle). Given that the average power draw can be reasonably assumed to remain constant until battery depletion, the expected battery life T_{life} can be predicted as $T_{\text{life}} = C_{\text{batt}}/P_{\text{loop}}$ being a function of the available battery capacity C_{batt}.

7. Concluding Remarks

In this paper a cooperative network architecture is proposed to emulate transmission and reception of data on a distributed network for tight closed-loop process control applications. A proprietary cooperative link-layer protocol has been developed on top of an existing IEEE 802.15.4 compliant PHY/MAC layer architecture to implement a virtual multiple-antenna array system. A multihop chain of

consecutive cooperative transmission sessions guarantees a robust two-way communication between the controller and the I/O sensor with a cycle time of 50 ms. The cooperative network protocol configuration imposes a substantial redefinition of conventional radio-planning methods. The required level of cooperative diversity for high quality control depends on the unstable properties that characterize the process in open loop. Despite the clear benefits of the proposed scheme, the experimental results highlighted a number of limitations that could be the target of future research: (i) compared to multihop architectures, the exploitation of cooperative diversity demands far more energy to enable the combining stage: the use of optimized batteries and/or harvesting from alternative sources of power represents promising solutions to improve device lifetime, another option is to enable event-driven control strategies to limit the channel use; (ii) a massive deployment of virtual controllers for the simultaneous control of multiple processes might cause spectrum overcrowding and autointerference: this suggests the adoption of advanced network-coding schemes to improve spectral efficiency; (iii) typical highly critical processes (e.g., motion control) cannot tolerate *any* interruption of feedback control, as this might result in costly losses for the plant operator: cable-replacing in highly critical loops is therefore not feasible for current low-power radio technology. Besides these limitations, experimental results clearly suggest that the use of the proposed architecture is a mandatory roadmap to enable cable-replacing in networked control systems.

Acknowledgments

This work has been partially supported by Saipem S.p.A, a subsidiary of Eni and has been performed in the framework of the European research project DIWINE (Dense Cooperative Wireless Cloud Network) under FP7 ICT Objective 1.1—The Network of the Future. The author would like to thank Dr. Sergio Guardiano, Professor Umberto Spagnolini, and Professor Vittorio Rampa for their fruitful discussions and helpful comments.

References

[1] D. Zuehlke, "Smart factory—towards a factory-of-things," *Annual Reviews in Control*, vol. 34, no. 1, pp. 129–138, 2010.

[2] P. Antsaklis and J. Baillieul, "Special issue on technology of networked control systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 5–8, 2007.

[3] European Commission, *Mixed Criticality Systems. Report From the Workshop on Mixed-Criticality Systems*, Brussels, Belgium, 2012.

[4] A. Willig, "Recent and emerging topics in wireless industrial communications: a selection," *IEEE Transactions on Industrial Informatics*, vol. 4, no. 2, pp. 102–124, 2008.

[5] S. Savazzi, U. Spagnolini, L. Goratti, D. Molteni, M. Latva-aho, and M. Nicoli, "Ultra-wide band sensor networks in oil and gas explorations," *IEEE Communications Magazine*, vol. 51, no. 4, pp. 142–153, 2013.

[6] F. De Pellegrini, D. Miorandi, S. Vitturi, and A. Zanella, "On the use of wireless networks at low level of factory automation systems," *IEEE Transactions on Industrial Informatics*, vol. 2, no. 2, pp. 129–143, 2006.

[7] J. Song, S. Han, A. K. Mok et al., "WirelessHART: applying wireless technology in real-time industrial process control," in *Proceedings of the 14th IEEE Real-Time and Embedded Technology and Applications Symposium (RTAS '08)*, pp. 377–386, April 2008.

[8] Standard ISA100.11a-2009, "Wireless systems for industrial automation: process control and related applications," ISA, July 2009.

[9] A. Sahai and S. Mitter, "The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link—part I: scalar systems," *IEEE Transactions on Information Theory*, vol. 52, no. 8, pp. 3369–3395, 2006.

[10] J. H. Braslavsky, R. H. Middleton, and J. S. Freudenberg, "Feedback stabilization over signal-to-noise ratio constrained channels," *IEEE Transactions on Automatic Control*, vol. 52, no. 8, pp. 1391–1403, 2007.

[11] Y. Ishido, K. Takaba, and D. E. Quevedo, "Stability analysis of networked control systems subject to packet-dropouts and finite-level quantization," *Systems & Control Letters*, vol. 60, no. 5, pp. 325–332, 2011.

[12] A. Scaglione, D. L. Goeckel, and J. N. Laneman, "Cooperative communications in mobile ad hoc networks," *IEEE Signal Processing Magazine*, vol. 23, no. 5, pp. 18–29, 2006.

[13] S. Savazzi and U. Spagnolini, "Energy aware power allocation strategies for multihop-cooperative transmission schemes," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 2, pp. 318–327, 2007.

[14] V. C. Gungor and G. P. Hancke, "Industrial wireless sensor networks: challenges, design principles, and technical approaches," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 10, pp. 4258–4265, 2009.

[15] S. Savazzi, S. Guardiano, and U. Spagnolini, "Wireless sensor network modeling and deployment challenges in oil and gas refinery plants," *International Journal of Distributed Sensor Networks*, vol. 2013, Article ID 155014, 17 pages, 2013.

[16] A. Bardella, N. Bui, A. Zanella, and M. Zorzi, "An experimental study on IEEE 802.15.4 multichannel transmission to improve RSSI—based service performance," in *Proceedings of the 4th International Workshop on Real-World Wireless Sensor Networks (REALWSN '10)*, pp. 154–161, Colombo, Sri Lanka, 2001.

[17] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. S. Sastry, "Foundations of control and estimation over lossy networks," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 163–187, 2007.

[18] O. C. Imer, S. Yüksel, and T. Başar, "Optimal control of LTI systems over unreliable communication links," *Automatica*, vol. 42, no. 9, pp. 1429–1439, 2006.

[19] J. Baillieul and P. J. Antsaklis, "Control and communication challenges in networked real-time systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 9–28, 2007.

[20] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 6, pp. 1089–1098, 2004.

[21] S. Savazzi and U. Spagnolini, "Cooperative fading regions for decode and forward relaying," *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 4908–4924, 2008.

[22] C. X. Wang, X. Hong, X.-H. Ge, X. Cheng, G. Zhang, and J. S. Thompson, "Cooperative MIMO channel models: a survey," *IEEE Communications Magazine*, vol. 48, no. 2, pp. 80–87, 2010.

[23] P. Castiglione, S. Savazzi, M. Nicoli, and T. Zemen, "Partner selection in indoor-to-outdoor cooperative networks: an experimental study," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 8, pp. 1–13, 2013.

[24] H. Shan, W. Zhuang, and Z. Wang, "Distributed cooperative MAC for multihop wireless networks," *IEEE Communications Magazine*, vol. 47, no. 2, pp. 126–133, 2009.

[25] H. Shan, H. T. Cheng, and W. Zhuang, "Cross-layer cooperative MAC protocol in distributed wireless networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2603–2615, 2011.

[26] P. Liu, Z. Tao, S. Narayanan, T. Korakis, and S. S. Panwar, "CoopMAC: a cooperative MAC for wireless LANs," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 2, pp. 340–353, 2007.

[27] B. Zhao and M. C. Valenti, "Practical relay networks: a generalization of hybrid-ARQ," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 1, pp. 7–18, 2005.

[28] H. Zhu and G. Cao, "rDCF: a relay-enabled medium access control protocol for wireless ad hoc networks," *IEEE Transactions on Mobile Computing*, vol. 5, no. 9, pp. 1201–1214, 2006.

[29] C. Shi, H. Zhao, S. Wang, J. Wei, and L. Zheng, "CAC-MAC: a cross-layer adaptive cooperative MAC for wireless ad-hoc networks," *International Journal of Distributed Sensor Networks*, vol. 2012, Article ID 155014, 9 pages, 2012.

[30] V. Nguyen and D. Brunelli, "Cooperative transmission range doubling with IEEE 802.15.4," in *Proceedings of IEEE International Conference on Communications*, pp. 126–130, June 2012.

[31] P. B. Sousa and L. L. Ferreira, "Hybrid wired/wireless profibus architectures: performance study based on simulation models," *EURASIP Journal on Wireless Communications and Networking*, vol. 2010, Article ID 845792, 25 pages, 2010.

[32] S. Ganeriwal, I. Tsigkogiannis, H. Shim, V. Tsiatsis, M. B. Srivastava, and D. Ganesan, "Estimating clock uncertainty for efficient duty-cycling in sensor networks," *IEEE/ACM Transactions on Networking*, vol. 17, no. 3, pp. 843–856, 2009.

[33] Datasheet CC2420, "2.4 GHz IEEE 802.15.4 ZigBee-ready RF Transceiver," March 2007.

[34] L. Angrisani, M. Bertocco, D. Fortin, and A. Sona, "Experimental study of coexistence issues between IEEE 802.11b and IEEE 802.15.4 wireless networks," *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 8, pp. 1514–1523, 2008.

[35] J. Lunze and D. Lehmann, "A state-feedback approach to event-based control," *Automatica*, vol. 46, no. 1, pp. 211–215, 2010.

[36] WirelessHART, "IEC 62591, System Engineering Guide," Revision 2, October 2010.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 920563, 5 pages
http://dx.doi.org/10.1155/2013/920563

Research Article

Immune Embedded Linux Core System with Multiple Sensors

Tao Gong[1,2,3]

[1] *College of Information S. & T., Donghua University, Shanghai 201620, China*
[2] *Engineering Research Center of Digitized Textile & Fashion Technology, Ministry of Education, Donghua University, Shanghai 201620, China*
[3] *Department of Computer Science, Purdue University, West Lafayette, IN 47907, USA*

Correspondence should be addressed to Tao Gong; taogong@dhu.edu.cn

Received 14 November 2012; Revised 28 March 2013; Accepted 28 March 2013

Academic Editor: Sabah Mohammed

The security of the embedded Linux core system is more important than before because this system is becoming more useful to solve the security problem of the computer system than the Windows operating system. It is firstly necessary to identify the normal state of the embedded Linux core system with a model. In this paper, the normal model of this normal system is represented with the space-time property set of all the components in this system. For example, the space property of a file in this system is the absolute pathname of this file, and the time property of this file is its last revision time. This immune embedded Linux core system uses multiple sensors to detect some important information, such as the disaster features. This antidisaster system will be tested on the ARM11 chips.

1. Introduction

As we know, an operating system such as Windows is often vulnerable to some viruses and attacks. This operating system must be updated with almost daily security patches, which cause new vulnerability due to the viruses or faults in the patches. It is true that the Linux is more secure than the Windows because the Linux is open at its source codes and can be enhanced in security by adding the security programs and customizing the source codes of the embedded Linux core system for special applications.

In fact, unknown viruses are difficult to be detected by traditional approaches [1] because the traditional security programs often detect the viruses with the feature matching of the viruses rather than those of the normal components. If these viruses are not prevented in the network, the network will be destroyed by the viruses in a short time [2].

To find the solution to the unknown viruses, we can seek inspiration from nature [3]. The human immune system really has advanced security mechanism to discriminate the selfs from the nonselfs and protect the body against the viruses, which are sometimes unknown [4]. The human immune system can detect the selfs first with immune tolerance to the selfs. We can build a normal model for the selfs to make the immune tolerance.

In this paper, by building an immune mechanism based on the normal model, an immune embedded Linux core system with multiple sensors is presented. In this system, the sensors are used to collect the data about the environment information of the disaster.

2. Related Work

In the histories of computers and networks, the viruses caused an increasing number of economic losses. On November 2nd, 1988, the Morris worms infected over 6000 Internet servers and the losses of these paralyzed servers were more than ten million dollars [5]. On July 19th, 2001, the CodeRed virus occurred and then caused the losses of more than 2 billion dollars. Moreover, the variants of this virus were more powerful and the variant CodeRed II caused the losses of more than 1.2 billion dollars. On September 18th, 2001, the Nimda virus occurred and then caused the losses of about 2.6 billion dollars.

To decrease the damage of the viruses, some network techniques were used to detect the viruses and to stop the

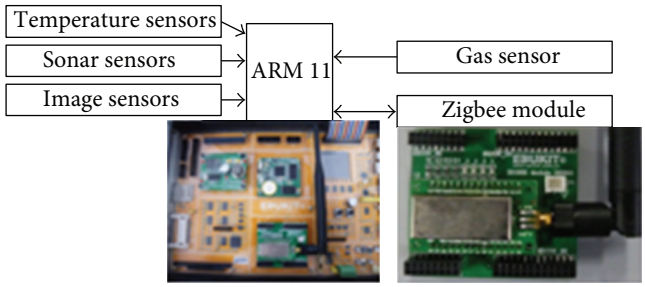


Figure 1: Frame of embedded Linux core system with ARM11 and multiple sensors.

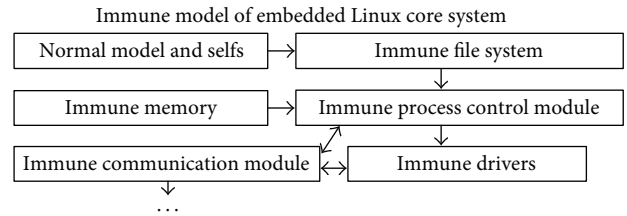


Figure 2: Immune model of embedded Linux core system.

spread [6–8]. The features of the viruses were analyzed and some models were built to describe the viruses and their spread. However, most of the traditional approaches cannot detect unknown viruses in real time, and the detection rates are not high enough. On the other hand, the variants of old viruses are continuously spread, and new viruses were increasingly designed. To overcome this bottleneck, it is necessary to design an immune embedded Linux core system based on the normal model of this core and more accurate self/nonself detection.

3. Embedded Linux Core System with Multiple Sensors

Suppose that this embedded Linux core system is composed of n_s antidisaster sensors, n_a ($n_s > n_a$) ARM11 chips, and n_a Zigbee modules, as shown in Figure 1. The antidisaster sensors include the temperature sensor, the sonar sensor, the image sensor (digit camera), and gas sensor and so on. The antidisaster sensors are used to collect all kinds of information about the disaster, and the Zigbee modules are used to transfer this information between an ARM11 node and another one.

This embedded Linux core system is built on the ARM11 chip, and the disaster environment information is collected by some sensors. The Zigbee modules make these ARM11 chips connected in a local wireless network. If a victim calls for help through an embedded system in Figure 1, the other embedded Linux core system will receive such information and activate the rescue procedure. If a software fault occurs in the embedded Linux core system, this system cannot be used to call for help or activate the rescue procedure. So it is very necessary and crucial to establish an intelligent mechanism to detect the faults and repair this damaged system in time. The immune system is one of the useful solutions for this problem.

4. Immune Model of Embedded Linux Core System

As we know, a complex standard Linux has too many components, so this operating system cannot be installed into the embedded system. Because the embedded system has limited memory to run programs, it does not need all the components of the standard Linux. So the standard Linux is trimmed into an embedded Linux, and the Linux core should be revised. The embedded Linux core system has necessary components such as files and directories. As shown in Figure 2, the immune model of this embedded Linux core system is built on its functional modules, which include file system, process control module, communication module, drivers, and memory management module.

5. Immune Algorithms for This Embedded Linux Core system

After analyzing the characteristics of the disaster samples of images, gas and temperatures, the real-coding affinity measure was used [9], and an immune algorithm was designed for recognizing the disaster on this embedded Linux core system with multiple sensors.

> *Input*: N training samples $x_1, x_2, \ldots, x_N$ and N_o disaster feature vectors, whose information is collected from the antidisaster sensors and denoted as o_α, $\alpha = 1, 2, \ldots, N_o$.
>
> *Output*: The recognition types of the disaster feature vectors $\{o_\alpha\}$.

Step 1. Initialize the parameters of the feature space, the feature-acquiring operator of each sensor, and the operator for searching the most similar sample.

Step 2. Build the feature space R^d with the feature vectors of the samples $x_1, x_2, \ldots, x_N$ and represent the antibody vectors Ab_m, $1 \le m \le N$.

Step 3. Collect the feature information from the antidisaster sensors, build the feature vector o_α with this feature information, and design the antigen Ag_α.

Step 4. Search the most similar sample of the object o_α with the real-coding clonal selection algorithm, by matching the antigen Ag_α with any antibody Ab_m, $m = 1, 2, \ldots, N$ according to that affinity measure.

Step 5. Calculate with uncertain reasoning and output the recognition type of the disaster feature vector o_α.

Step 6. Transfer the recognition result of the disaster feature vector o_α to other ARM11 chips.

```
temperature4_warning!
Ok!The value of combustible gas is:     1.0145
temperature1_warning!
temperature2_warning!
temperature3_warning!
temperature4_warning!
Ok!The value of combustible gas is:     1.0137
temperature1_warning!
temperature2_warning!
temperature3_warning!
temperature4_warning!
Ok!The value of combustible gas is:     1.0105
```

Figure 3: Abnormal temperature detection of the antidisaster sensor on the embedded Linux core system.

Moreover, another immune algorithm was also designed to build an initial normal model of the embedded Linux core system.

Step 1. Initialize the self-database.

Step 2. Make the backup system of the initial normal embedded Linux core system S and provide the root directory of the backup system.

Step 3. Read the root directory of the system S and search all the files in the root directory of this system.

Step 4. If there is at least one unread file or unread subdirectory in the current directory, then select one and read its absolute pathname and its last revision time; otherwise, set the ending condition satisfied and go to Step 7.

Step 5. Append the absolute pathname and the last revision time of this file or subdirectory as an immune record object into the self database.

Step 6. If this object is a file, then close the pointer of this file; otherwise, recursively build the normal model of the subsystem for this subdirectory.

Step 7. If the ending condition is satisfied, then end; otherwise, go to Step 3.

Based on the initial normal model and the rules for transforming a normal state to another one, the normal model of the immune embedded Linux core system can be updated dynamically.

With this normal model and the rules for transforming the normal states, the viruses and software faults can be detected by detecting the normal components (selfs) first in this embedded Linux core system. This detection algorithm is based on the self/nonself discrimination via the normal model.

Step 1. Initialize the infection database.

Step 2. Read the root directory of the system S and search all the files in the root directory of this system.

Step 3. If there is at least one unread file or unread subdirectory in the current directory, then select one and read its absolute pathname and its last revision time; otherwise, set the ending condition satisfied and go to Step 7. If the absolute pathname and last revision time of a file or directory are not matched with those of any record in the self database, then this file or directory is not a self; that is, this is a non-self, which may be a virus or software fault. Then this virus or software fault will be recognized and eliminated.

Step 4. Append the absolute pathname and backup pathname of the infected file or the infected subdirectory into the infection database.

Step 5. Recognize the viruses and software faults with feature matching and immune learning.

Step 6. Eliminate the viruses and software faults and then repair the damaged system.

Step 7. If the ending condition is satisfied, then end; otherwise, go to Step 3.

6. Experimental Results

Some experiments were conducted to test the antidisaster detection of the embedded Linux core system with antidisaster sensors. All the antidisaster sensors were tested on the ARM11 chips, and the data were analyzed on the PC machine, as shown in Figure 3.

From the monitor data of the antidisaster sensor, the temperature curve with the abnormal temperature is shown in Figure 4, and the gas voltage curve with the harmful gas from the sensor is shown in Figure 5.

The disaster recognition was built on the data from the multiple antidisaster sensors and the immune algorithm with clonal selection and real-coding affinity measure. The results of the recognition and the self/non-self detection were transferred between one ARM chip and another one, as shown in Figure 6. The received data showed the data source first and then the transferred data. When the victims need help in a disaster, the embedded Linux core system with the multiple antidisaster sensors were used to call for help and share the information about the disaster.

In these experiments, the immune Linux core was compiled with some ordinary compiling methods, and the step for compiling immune programs was started after building the normal model of this embedded Linux core system. After these immune programs were compiled, the files Makefile and Konfig were regenerated. The mirror of the Linux core used the zImage mirror, and this mirror was downloaded

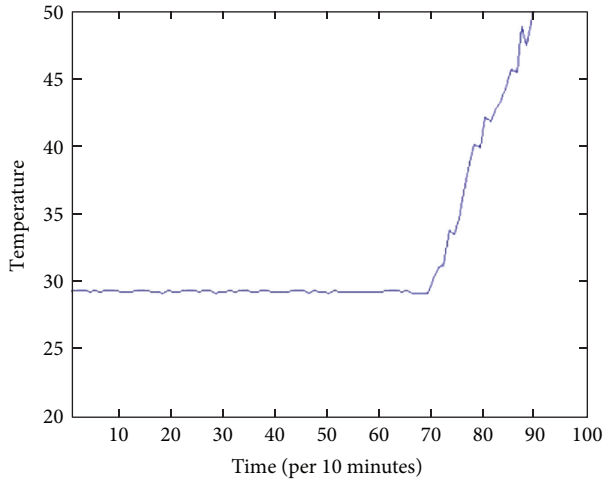


Figure 4: Temperature curve with abnormal temperature from the sensor.

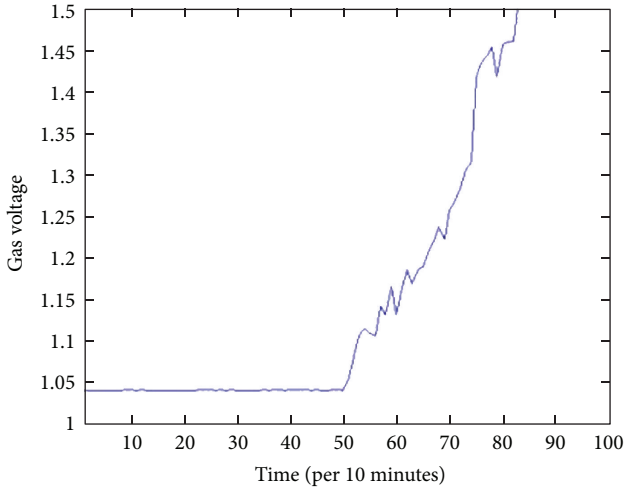


Figure 5: Gas voltage curve with harmful gas from the sensor.

```
T1 : 28.9375C

T2 : 28.9375C

T3 : 28.9375C

T4 : 28.9375C

Ok!The value of combustible gas is:     1.0645

T1 : 28.9375C

eceive data from :  0001
he data is :  123456789
T2 : 28.8750C

T3 : 28.8750C

T4 : 28.8750C

Ok!The value of combustible gas is:     1.0621
```

Figure 6: Data transferring between the ARM11 chips via the Zigbee modules.

```
Now building Normal Model…/home/core/rootfs/var/lib
Epoch Time is 1339382533
Now building Normal Model…/home/core/rootfs/home
Epoch Time is 1339382466
Now building Normal Model…/home/core/rootfs/opt
Epoch Time is 1345520051
Now building Normal Model…/home/core/rootfs/sys
Epoch Time is 1339382466

Normal Model has been built successfully!
File count is 672.
```

Figure 7: Building the normal model of this embedded Linux core system.

```
Now, start self/nonself detection...
Nonselfs detected are printed on the screen.

Nonself detected: /home/core/rootfs/etc/init.d
Nonself detected: /home/core/rootfs/etc/init.d/rcS~
Nonself detected: /home/core/rootfs/etc/init.d/rcS
Self/Nonself detection results are:
Selfs: 669, Nonselfs: 3
```

Figure 8: Self/nonself detection of Linux core system with immune programs.

```
Start system self-repairing...
cp: /home/core/backup/rootfs/etc /home/core/rootfs/etc
cp: omitting directory ‘/home/core/backup/rootfs/etc’
cp: /home/core/backup/rootfs/etc/group~ /home/core/rootfs/etc/group~
cp: /home/core/backup/rootfs/etc/group /home/core/rootfs/etc/group
cp: /home/core/backup/rootfs/etc/init.d /home/core/rootfs/etc/init.d
cp: omitting directory ‘/home/core/backup/rootfs/etc/init.d’
cp: /home/core/backup/rootfs/etc/init.d/rcS~ /home/core/rootfs/etc/init.d/rcS~
cp: /home/core/backup/rootfs/etc/init.d/rcS /home/core/rootfs/etc/init.d/rcS
cp: /home/core/backup/rootfs/etc/passwd /home/core/rootfs/etc/passwd
cp: /home/core/backup/rootfs/etc/passwd~ /home/core/rootfs/etc/passwd~
```

Figure 9: System repairing based on the normal model and immune programs.

online into the development board to test. These immune programs were tested to build a normal model for this immune Linux core system, as shown in Figure 7.

After this embedded Linux core system was changed by some nonselfs, the nonselfs were detected by the immune programs, as shown in Figure 8.

All the nonselfs were eliminated, and then the damaged Linux core system was repaired with the normal model and the immune programs, as shown in Figure 9.

7. Conclusions

With the increasing requirements about the security of the operation system and network, the embedded Linux core system is more crucial to the security applications. Inspired by the nature and beyond the traditional security approaches, the model of immune computation becomes a new effective tool to enhance the security of the Linux operatiing system from the core to some applications. In this research, the immunization mechanism is embedded into the Linux core

system to keep its functional modules, such as the file system, immune, and secure.

Acknowledgments

This work was supported in part by grants from the National Natural Science Foundation of China (61271114), the Natural Science Foundation of Shanghai (08ZR1400400 and 11ZR1401300), the Shanghai Postgraduate Education Creative Plan (SHGS-KC-2012OO3), the Shanghai Educational Development Foundation (2007CG42 and 12CG35), and the Fundamental Research Funds for the Central Universities at Donghua University (13D110414). The author thanks his graduate student Changxing Du for the help in programming.

References

[1] O. Sukwong, H. S. Kim, and J. C. Hoe, "Commercial antivirus software effectiveness: an empirical study," *IEEE Computer*, vol. 44, no. 3, Article ID 5506074, pp. 63–70, 2011.

[2] J. Balthrop, S. Forrest, M. E. J. Newman, and M. M. Williamson, "Technological networks and the spread of computer viruses," *Science*, vol. 304, no. 5670, pp. 527–529, 2004.

[3] T. Gong and C. X. Du, "Immune computation of secure embedded linux core against viruses and software faults," *International Journal of Security and its Applications*, vol. 6, no. 2, pp. 167–172, 2012.

[4] R. Medzhitov and C. A. Janeway, "Decoding the patterns of self and nonself by the innate immune system," *Science*, vol. 296, no. 5566, pp. 298–300, 2002.

[5] H. Orman, "The morris worm: a fifteen-year perspective," *IEEE Security and Privacy*, vol. 1, no. 5, pp. 35–43, 2003.

[6] E. Levy, "Worm propagation and generic attacks," *IEEE Security and Privacy*, vol. 3, no. 2, pp. 63–65, 2005.

[7] B. Madhusudan and J. W. Lockwood, "A hardware-accelerated system for real-time worm detection," *IEEE Micro*, vol. 25, no. 1, pp. 60–69, 2005.

[8] I. Arce and E. Levy, "An analysis of the slapper worm," *IEEE Security and Privacy*, vol. 1, no. 1, pp. 82–87, 2003.

[9] T. Gong, "High-precision immune computation for secure face recognition," *International Journal of Security and Its Applications*, vol. 6, no. 2, pp. 293–298, 2012.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 404168, 15 pages
http://dx.doi.org/10.1155/2013/404168

Research Article

Novel MAC Protocol and Middleware Designs for Wearable Sensor-Based Systems for Health Monitoring

Kyeong Hur,[1] Won-Sung Sohn,[1] Jae-Kyung Kim,[1] and YangSun Lee[2]

[1] *Department of Computer Education, Gyeongin National University of Education, Gyesan-Dong, San 59-12, 45 Gyodae-Gil, Gyeyang-Gu, Incheon 407-753, Republic of Korea*
[2] *Department of Computer Engineering, Seokyeong University, 16-1 Jungneung-Dong, Sungbuk-Ku, Seoul 136-704, Republic of Korea*

Correspondence should be addressed to Won-Sung Sohn; sohnws@ginue.ac.kr

Received 16 November 2012; Accepted 11 December 2012

Academic Editor: Sabah Mohammed

We propose a middleware platform built on wireless USB (WUSB) over wireless body area networks (WBAN) hierarchical protocol for wearable health-monitoring systems (WHMS). The proposed middleware platform is composed of time-synchronization and localization solutions. It is executed on the basis of WUSB over WBAN protocol at each wearable sensor node comprising the WHMS. In the platform, firstly, the time-synchronization middleware is executed. After that, a WHMS host calculates the location of a receiving sensor node by using the difference between the times at which the sensor node received different WBAN beacon frames sent from the WHMS host. The WHMS host interprets the status and motion of the wearable body-sensor objects.

1. Introduction

Wearable health-monitoring systems (WHMS) have drawn a lot of attention from the research community and the industry during the last decade as it is pointed out by the numerous and yearly increasing corresponding research and development efforts [1]. To address this demand, a variety of system prototypes and commercial products have been produced in the course of recent years, which aim at providing real-time feedback information about one's health condition, either to the user himself, to a medical center, or straight to a supervising professional physician, while being able to alert the individual in case of possible imminent health-threatening conditions. In addition to that, WHMS constitute a new means to address the issues of managing and monitoring chronic diseases, elderly people, postoperative rehabilitation patients, and persons with special abilities [1, 2].

Wearable systems for health monitoring may comprise various types of miniature sensors, wearable or even implantable [1]. These biosensors are capable of measuring significant physiological parameters like heart rate, blood pressure, body and skin temperature, oxygen saturation, respiration rate, electrocardiogram, and so forth. The obtained measurements are communicated either via a wireless or a wired link to a central node, for example, a personal digital assistant (PDA) or a microcontroller board, which may then in turn display the according information on a user interface or transmit the aggregated vital signs to a medical center. The previous illustrates the fact that a wearable medical system may encompass a wide variety of components: sensors, wearable materials, smart textiles, actuators, power supplies, wireless communication modules and links, control and processing units, interface for the user, software, and advanced algorithms for data extracting and decision making.

The wireless body area network (WBAN) is a promising technology that can revolutionize next-generation healthcare applications. Developing a unifying WBAN standard that addresses the core set of technical requirements is the quintessential step to unleash the full potential of WBAN and is currently under discussion in the IEEE 802.15.6 Task Group [2]. The IEEE 802.15.6 WBAN standard is used in or around a body [2, 3]. It is designed to serve advanced medical and entertainment options enabled by this standard. It will allow medical equipment manufacturers and consumer electronics manufacturers to have small, power-efficient, and inexpensive solutions to be implemented for a wide range of devices. This standard considers effects on portable antennas due to the presence of a person (varying with male, female,

skinny, heavy, etc.), radiation pattern shaping to minimize specific absorption rate (SAR) into the body, and changes in characteristics as a result of the user motions.

WiMedia Alliance is developing the specifications of the PHY, MAC, and convergence layers for Ultrawideband (UWB) systems with participation from more than 170 companies. Also, it has been promoting the standardization and adaptation of UWB for high rate-wireless personal area network (HR-WPAN) that enables the multimedia and high-speed data communication [4, 5]. WiMedia Alliance has completed the specification of WiMedia distributed-MAC (D-MAC), and this enables various applications, such as wireless USB (WUSB) [4], Wireless 1394, and Wireless IP, to operate on WiMedia D-MAC. The WiMedia D-MAC supports a distributed-MAC approach. In contrast to IEEE 802.15.3, D-MAC makes all devices have the same functionality, and networks are self-organized and provide devices with functions such as access to the medium, channel allocation to devices, data transmission, quality of service, and synchronization in a distributed manner [5].

The term middleware refers to the software layer located between the OS and the application layer. Middleware can be further classified based on "submiddleware" functionalities including time synchronization, location, battery power, and network functionalities. As in all distributed systems, time synchronization is very important in a sensor network since the design of many protocols and the implementation of applications require precise timing information. In forming an energy-efficient radio schedule, conducting in-network processing such as data fusion, performing RF ranging, and tracking a certain object; for example, time synchronization is certainly required.

Time-synchronization problem has been investigated thoroughly in Internet and LANs. Complex protocols such as NTP [6] have been developed to keep the Internet's clocks ticking in phase. However, the peculiar constraints of sensors and adhoc network topologies of wireless sensor networks (WSNs) impose specific requirements on protocol design of time synchronization for WSN applications. First, since the amount of energy available to battery-powered sensors is quite limited, time synchronization must be implemented in an energy-efficient way. Second, the small size of a sensor node imposes restrictions on computational power and storage space. Therefore, traditional synchronization schemes such as NTP and GPS are not suitable for WSNs because of complexity and energy issues, cost efficiency, limited size, and so on [7].

In this paper, we propose a middleware platform built on wireless USB (WUSB) over wireless body area networks (WBANs) hierarchical protocol for wearable health-monitoring systems (WHMS). The proposed middleware platform is composed of time-synchronization and localization solutions. It is executed on the basis of WUSB over WBAN protocol at each sensor node comprising the WHMS. In the platform, firstly, the time-synchronization middleware is executed. After that, a WHMS host calculates the location of a receiving sensor node by using the difference between the times at which the sensor node received different WBAN beacon frames sent from the WHMS host. The WHMS host interprets the status and motion of the attached body-sensor objects.

2. Features of IEEE 802.15.6 WBAN Protocol

To provide or support time-referenced allocations in its wireless body area network (WBAN), a hub shall establish a time base as specified in [3] which divides the time axis into beacon periods (superframes) regardless of whether it is to transmit beacons. In such cases, the hub shall transmit a beacon in each beacon period, except in inactive superframes, unless prohibited by regulations such as imposed in MICS band. The hub may shift (rotate) its beacon transmission time from one offset from the start of current beacon period to another offset from the start of the next beacon period, thereby shifting the time reference for all scheduled allocations, to prevent large-scale repeated transmission collisions between its WBAN and neighbor WBANs [3].

In cases where a hub is not to provide or support time-referenced allocations in its WBAN, it may operate with establishing neither a time base nor superframe boundaries and hence without transmitting beacons at all. Equivalently, a hub shall operate in beacon mode transmitting a beacon in every beacon period other than in inactive superframes to enable time-referenced allocations, unless regulations as applicable in the MICS band disallow beacon transmission. In the latter case, a hub shall operate in nonbeacon mode transmitting no beacons, with superframe and allocation slot boundaries established if access to the medium in its WBAN involves time referencing, or without superframe or allocation slot boundaries if access to the medium in its WBAN involves no time referencing. In the beacon mode, a hub shall divide each active beacon period into applicable access phases as illustrated in Figure 1. The hub may announce some superframes (beacon periods) as inactive superframes where it transmits no beacons and provides no access phases if there are no allocation intervals scheduled in those superframes [3].

The hub shall place the access phases—exclusive access phase 1 (EAP1), random access phase 1 (RAP1), type I/II access phase, exclusive access phase 2 (EAP2), random access phase 2 (RAP2), type I/II access phase, and contention access phase—in the order stated and shown above. The hub may set to zero the length of any of these access phases but shall not have RAP1 shorter than the guaranteed minimum length communicated in connection assignment frames sent to nodes that are still connected with it. To provide a nonzero length CAP, the hub shall transmit a preceding B2 frame [3].

A type I/II access phase is either a type I or type II access phase as described below but not both. The two type I/II access phases may be both of type I, both of type II, or one of type I and the other of type II. If one is of type I and the other is of type II, either one may appear before the other. The hub may schedule uplink allocation intervals, downlink allocation intervals, and bilink allocation intervals anywhere in a type I access phase. It may improvise type I and type II polled allocation intervals as well as posted allocation intervals anywhere outside the scheduled allocation intervals

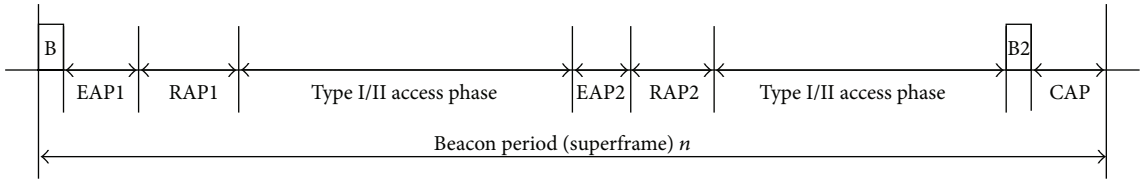


Figure 1: Layout of access phases in a beacon period (superframe) for beacon mode.

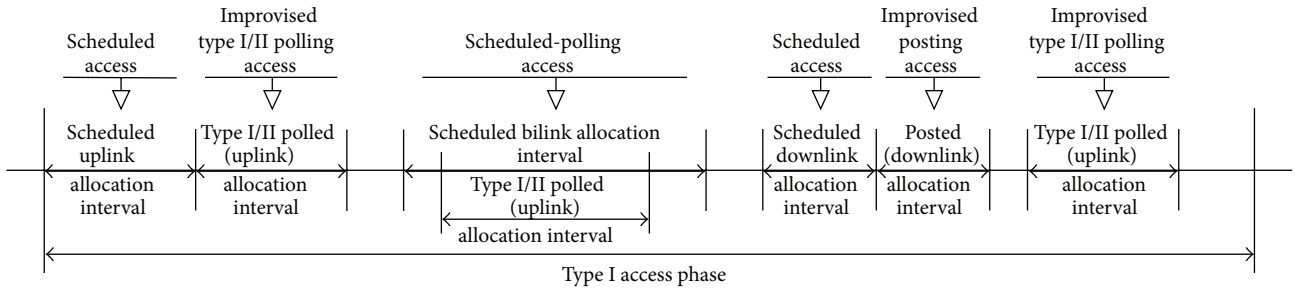


Figure 2: Allocation intervals and access methods permitted in a type I access phase.

in the type I access phase. It may also provide type I and type II polled allocation intervals within scheduled bilink allocation intervals in a type I access phase.

A type I polled allocation is conveyed in terms of its time duration (the maximum time the polled node may use for its frame transactions in the allocation). And a type II polled allocation is conveyed in terms of a frame count (the maximum number of frames the polled node may transmit in the allocation). These allocation intervals along with the corresponding access methods by which they are obtained are illustrated in Figure 2.

The hub may schedule bilink allocation intervals and delayed bilink allocation intervals anywhere in a type II access phase, except that it shall not schedule any bilink allocation intervals after delayed bilink allocation intervals in the same type II access phase. It may improvise type II, but not type I, polled allocation intervals as well as posted allocation intervals anywhere outside the scheduled bilink allocation intervals and delayed bilink allocation intervals in the type II access phase. It may also provide type II, but not type I, polled allocation intervals within scheduled bilink allocation intervals and delayed bilink allocation intervals. These allocation intervals along with the corresponding access methods by which they are obtained are illustrated in Figure 3 [3].

In exclusive access phase 1(EAP1), random access phase 1 (RAP1), exclusive access phase 2 (EAP2), random access phase 2 (RAP2), and contention access phase (CAP), allocations may only be contended allocations, which are nonreoccurring time intervals valid per instance of access. The access method for obtaining the contended allocations shall be CSMA/CA if pRandonAccess is set to CSMA/CA, or slotted Aloha access if pRandonAccess is set to slotted Aloha.

A hub or nodes may obtain contended allocations in EAP1 and EAP1 if it needs to send data type frames of the highest user priority (i.e., containing an emergency or medical event report). The hub may obtain such a contended allocation pSIFS after the start of EAP1 or EAP2 without actually performing the CSMA/CA or slotted Aloha access procedure. Only nodes may obtain contended allocations in RAP1, RAP2, and CAP to send management or data type frames [3].

To send data type frames of the highest user priority based on CSMA/CA, a hub or a node may treat the combined EAP1 and RAP1 as a single EAP1, and the combined EAP2 and RAP2 as a single EAP2, so as to allow continual invocation of CSMA/CA and improve channel utilization. To send data type frames of the highest user priority based on slotted Aloha access, a hub or a node may treat RAP1 as another EAP1 but not a continuation of EAP1, and RAP2 as another EAP2 but not a continuation of EAP2, due to the time-slotted attribute of slotted Aloha access.

3. Data Flow in WUSB Protocol

WUSB is the technology-merged USB with UWB based on success of wired USB, and it can apply to WPAN applications as well as PAN applications like wired USB. Because WUSB specification has defined high-speed connection between a host and a device for the compatibility with USB 2.0 specification, it can be adapted easily for wired USB applications. WUSB connects WUSB devices with the WUSB host using a "hub and spoke" model [4]. The WUSB host is the "hub" at the center, and each device sits at the end of a "spoke." Each spoke is a point-to-point connection between the host and the device. Like this, the network formed by one host and several devices is referred to as the WUSB cluster.

WUSB hosts can support up to 127 devices and because WUSB does not have physical ports there is no need, nor any definition provided, for hub devices to provide port expansion. There is only one host in any WUSB cluster. And that host performs to transmit/receive a data with devices in the WUSB cluster. Also, it schedules the exchange of data

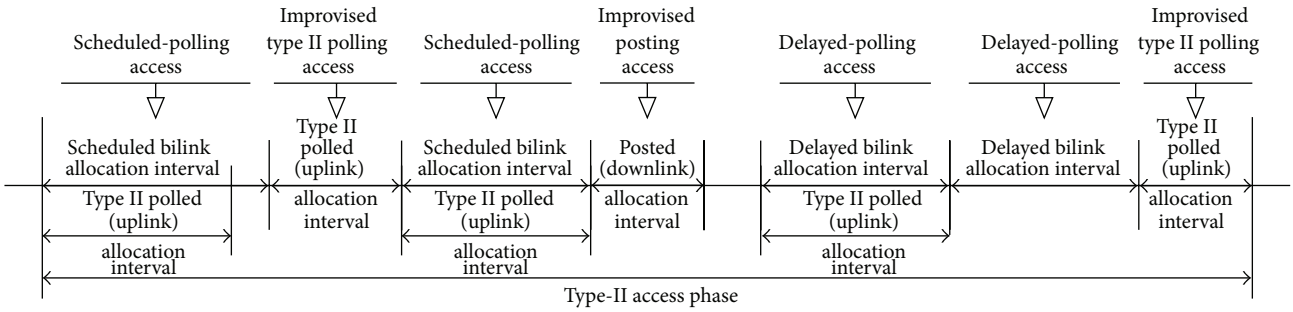


Figure 3: Allocation intervals and access methods permitted in a type II access phase.

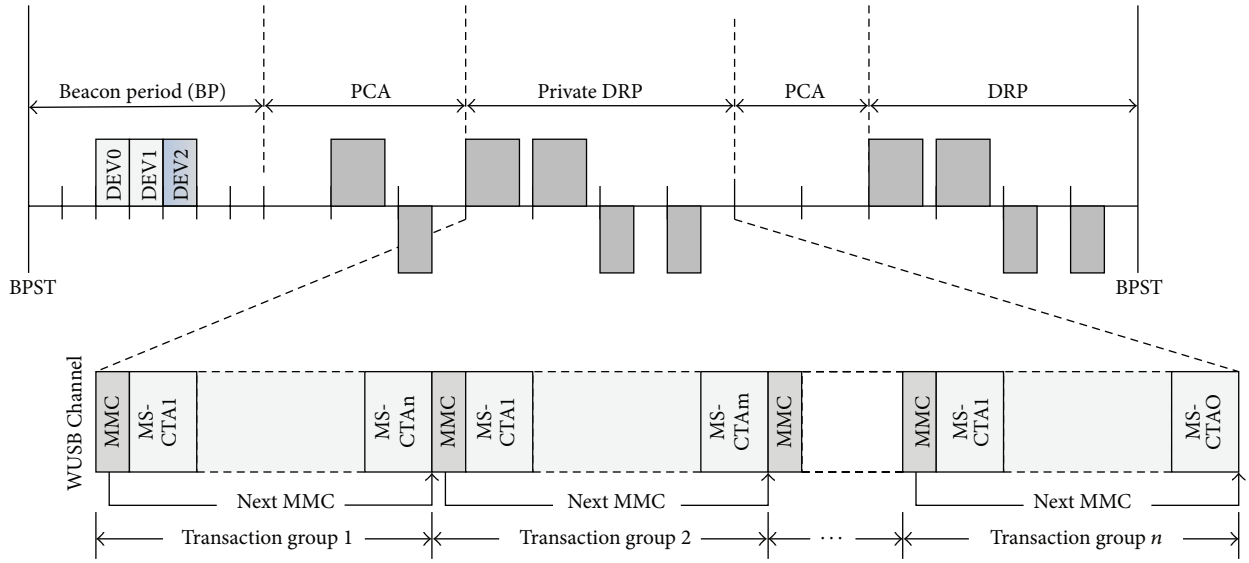


Figure 4: The example of the data exchange between WUSB devices through WiMedia D-MAC.

between WUSB host and WUSB devices and allocates time slots and channel bandwidths to WUSB devices in its own cluster. Because each WUSB cluster can overlap with the other at minimum interference, it can coexist with several WUSB clusters within the same communication environment. The distributed nature of D-MAC protocol can provide full mobility support and achieves scalable, fault-tolerant medium access method [4]. Thus, WUSB protocol based on WiMedia D-MAC can provide full mobility support.

WUSB defines a WUSB Channel which is encapsulated within a WiMedia MAC superframe via private DRP reservation blocks. The WUSB Channel is a continuous sequ-ence of linked application-specific control packets, called microscheduled management commands (MMCs), which are transmitted by the host within the private DRP reservation blocks. Figure 4 shows the relationship between WiMedia MAC and WUSB.

The microscheduled management command (MMC) is the fundamental element of the Wireless USB protocol. MMCs are used to help devices discover information about a WUSB cluster, notify their intentions, manage power, and schedule data transmissions efficiently to attain very high throughputs. The general structure of an MMC Control packet is showed in Figure 5 and detailed in Table 1.

A WUSB Channel consists of a continuous sequence of MMC transmissions from the host. The linked stream of MMCs is used primarily to dynamically schedule channel time for data communications between host applications and WUSB endpoints. An MMC specifies the sequence of microscheduled channel time allocations (MS-CTAs) up to the next MMC within a reservation instance or to the end of a reservation instance. It may be followed by another MMC without the existence of MS-CTAs between the two MMCs. In this case, the MMC is only used to convey command and control information. The channel time between two MMCs may also be idle time, where no MS-CTAs are scheduled. The MS-CTAs within a reservation instance can only be used by the devices that are members of the associated WUSB cluster. The direction of transmission and the use of each MS-CTA are fully declared in each MMC instance. An MMC can declare an MS-CTA during any channel time following the MMC.

An MMC contains the information elements necessary to identify the WUSB Channel or declare any MS-CTAs or

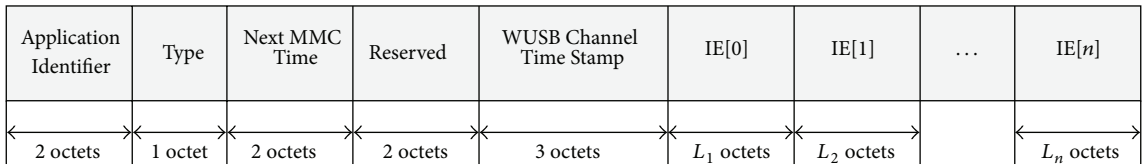


Figure 5: The general structure of an MMC Control packet.

Table 1: Detailed field description of MMC Control packet.

Field	WUSB equivalents
Application Identifier	Wireless USB: 0100H
Type	MMC command type: 01H
Next MMC Time	The number of microseconds from the beginning of this field indicates the value of the next MMC packet
Reserved	This field is reserved and should be set to "zero"
WUSB Channel Time Stamp	Timestamp provided by the host based on a free running timer in the host. The value in this field indicates the value of the host free running clock when MMC transmission starts. This timestamp is formatted into two fields as follows: (i) 0~6 bits: microsecond count. The microsecond count rolls over after 125 microseconds. Each time it rolls over the 1/8th millisecond count is incremented (ii) 7~23 bits: 1/8th millisecond count. This counter increments every time the microsecond counter wraps
IE $[0 - n]$	Array of information elements. There must be at least one IE

other information elements that are used for command and control. The MMC is a broadcast control packet that is for receipt only by devices that are members of the WUSB cluster. The host must use the broadcast cluster ID value in the DestAddr field of an MMC packet's MAC, WiMedia MAC header. This technique identifies this packet transmission as a broadcast targeting all devices in a WUSB cluster and avoids potential confusion at non-WUSB devices in listening range of the host. The MMC payload must be encapsulated within a WiMedia MAC secure packet; however, its data payload is transmitted in plain text, thus using the security encapsulation for authentication purpose only.

A host is required to implement the WiMedia MAC protocol and to establish and maintain WUSB Channels by allocating sequences of private DRP Reservation in WiMedia MAC. A device may implement the full WiMedia MAC protocol; however, it is nominally only required to implement the WUSB protocol which operates within the WUSB Channel.

As mentioned above, a WUSB host and WUSB devices must include a DRP IE in their beacon frames to protect the WUSB Channel. When a WUSB host becomes active, it must choose a PHY channel to operate as a WUSB Channel. Once the host is beaconing, it then establishes a WUSB Channel by DRP reservation for WUSB data communications. In this case, WUSB host is the DRP Reservation Owner, and WUSB device is the DRP Reservation Target. Thus, WUSB device must be able to determine which MASs are available for communication with the WUSB host. Since there are various devices in the WUSB cluster, the WUSB device identifies the WUSB host's DRP IE based on the following keys.

(i) Reservation type field is private.

(ii) Stream index field has the value of the WUSB Channel's stream index.

(iii) Owner DevAddr field is set to the WUSB Channel's broadcast cluster ID.

The WUSB device identifies a cluster member's DRP IE based on the following keys.

(i) Reservation Type field is Private.

(ii) Stream Index field has the value of the WUSB Channel's stream index.

(iii) Owner DevAddr field is set to the WUSB host's DevAddr.

Figure 6 shows the current operation of DRP reservation in WUSB protocol. A WUSB host uses the GetStatus (MAS Availability) request shown in Figure 7 to retrieve a device's MAS Availability information. A WUSB device that receives the GetStatus (MAS Availability) request from the WUSB host accumulates the information from its neighbors' beacon about available MASs. Then, the WUSB device responds to the GetStatus (MAS Availability) request through the bmMASAvailability field in the GetSatus request. In the format of the GetStatus request, bmMASAvailability field is a 256-bit map, where each bit location corresponds to an MAS slot in the WiMedia D-MAC Layer superframe. A 1 B in a bit location means that the device is available for a reservation in the corresponding MAS slot. A 0 B indicates that the device is not available.

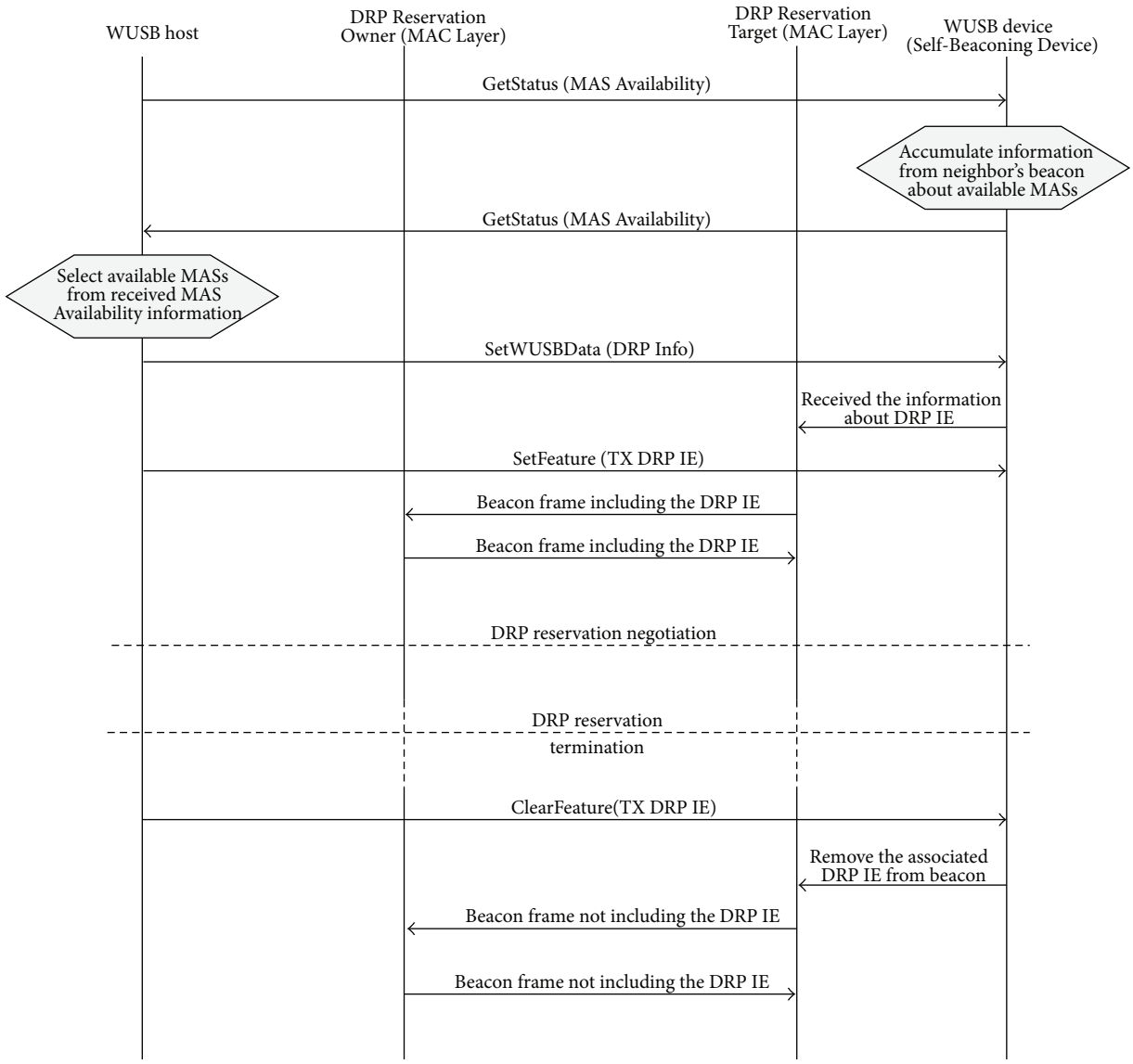


Figure 6: The current operation of DRP reservation in WUSB protocol.

bmRequest type (=10000000B)	bRequest (=GET_STATUS)	wValue (=0)	wIndex (=0004H)	wLength	bmMASAvailability (only returned by WUSB device)
1 octet	1 octet	2 octets	2 octets	2 octets	32 octets

Figure 7: The format of the GetStatus (MAS Availability) request.

After the WUSB host that receives the WUSB device's response selects available MASs, it transmits a SetWUSBData (DRP Info) request. The SetWUSBData (DRP Info) is used to construct the DRP IE that the WUSB device transmits in its beacon. If the WUSB device does not have an existing DRP IE for this Wireless USB Channel, it simply adds the received DRP IE to its beacon. If the device has an existing DRP IE for this Wireless USB Channel, then it must replace the existing DRP IE with the new DRP IE provided in this command payload. Figure 8 shows the format of SetWUSBData (DRP Info) request.

The values of bmAttributes field are used to construct the DRP control field of a DRP IE. The Conflict Tie-Breaker bit is set to a random value of zero or one when a reservation request is made, and it is used for DRP conflict resolution. DRPIEData field is the DRP allocation blocks that must be

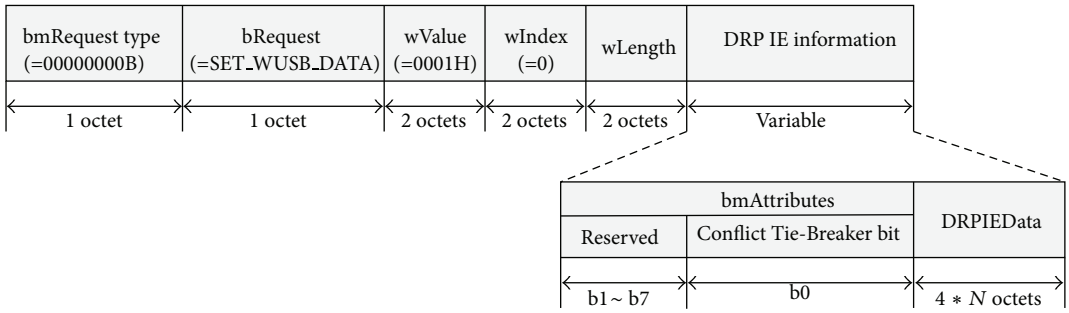


Figure 8: The format of the SetWUSBData (DRP Info) request.

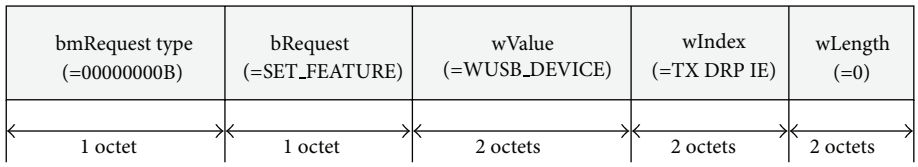


Figure 9: The format of the SetFeature (TX_DRPIE) request.

included in the DRP IE transmitted by the device. If the WUSB host transmits the SetFeature (TX_DRPIE) request to the WUSB device, the WUSB device starts the transmission of its beacon including DRP IE set to the values in DRP IE information field. Figure 9 shows the format of SetFeature (TX_DRPIE) request.

If the WUSB host wants to terminate the DRP reservations with the WUSB device, it transmits the ClearFeature (TX_DRPIE) request. On receipt of this request from the WUSB host, the WUSB device removes the associated DRP IE from its own beacon. Figure 10 shows the format of ClearFeature (TX_DRPIE) request.

4. WUSB over WBAN Architecture for WHMS

WBAN slave devices which have received beacon from WBAN host schedule their receiving and transmitting operations according to the information delivered by the beacon. IEEE 802.15.6 WBAN superframe begins with a beacon period (BP) in which the WBAN hub performing the WUSB host's role sends the beacon. The data transmission period in each superframe is divided into the exclusive access phase 1 (EAP1), random access phase 1 (RAP1), type I/II access phase, EAP2, RAP2, type I/II access phase, and contention access phase (CAP) periods. The EAP1 and EAP2 periods are assigned through contention to data traffic with higher priorities. Further, the RAP1, RAP2, and CAP periods are assigned through contention to data traffic with lower priorities.

In the WBAN beacon frame of Figure 11, the Sender Address field is set to the IEEE MAC address of the WBAN hub sending the current beacon. The Beacon Period Length field is set to the length of the current beacon period (superframe) in units of allocation slots. It is set to 0 to encode a value of 256 allocation slots. The random access phase 1 (RAP1) and RAP2 start fields are set to the number of the allocation slot that starts RAP1 and RAP2, respectively. The Beacon Shifting Sequence, Channel Hopping State, Next Channel Hop, and Inactive Duration fields are used for interference avoidance in WBAN wireless channel environment.

The IEEE 802.15.6 WBAN MAC systems have several MAC Capability options. Figure 12 shows current WBAN MAC Capability format standard. To successfully set up wireless communication links for WHMS, secure WUSB Channels should be encapsulated privately within a WBAN superframe. Furthermore, the WUSB Channel allocated privately in the IEEE 802.15.6 WBAN MAC enables the MMC scheduling between WUSB host and its several peripheral devices. In the current IEEE 802.15.6 WBAN standard, there are no available periods allocated for other heterogeneous networks.

In Figure 13 for a new WBAN MAC Capability format, a Private Period Allocation field is made by using one bit in the reserved bits in Figure 12. The Private Period Allocation field is set to one if the WBAN MAC supports private WUSB Channel allocations or set to 0 otherwise. If the Private Period Allocation field is set to 1, the RAP2 length field indicates the private WUSB Channel Length for MMC scheduling. Therefore, as shown in Figure 14, after receiving beacon from the WBAN host, non-WUSB WBAN slave devices enter into sleep mode during the RAP2 period. On the contrary, the WUSB/WBAN slave devices enter into active mode during the RAP2 period and they enter into sleep mode during the other periods. The RAP2 Length is determined by the WUSB/WBAN host according to priorities between WUSB transactions and WBAN traffic.

Figure 15 shows a new IEEE 802.15.6 superframe structure supporting the WUSB private channel allocation. The WUSB/WBAN host should transmit WUSB data without interference with WBAN data when a request for WUSB data transmissions occurs in the WUSB cluster. For this

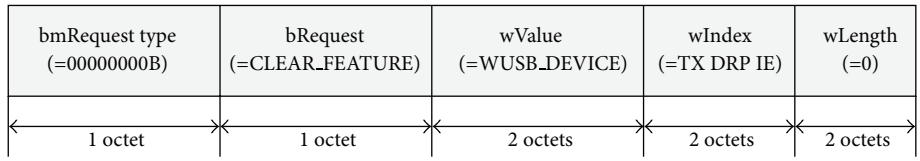


Figure 10: The format of the ClearFeature (TX_DRPIE) request.

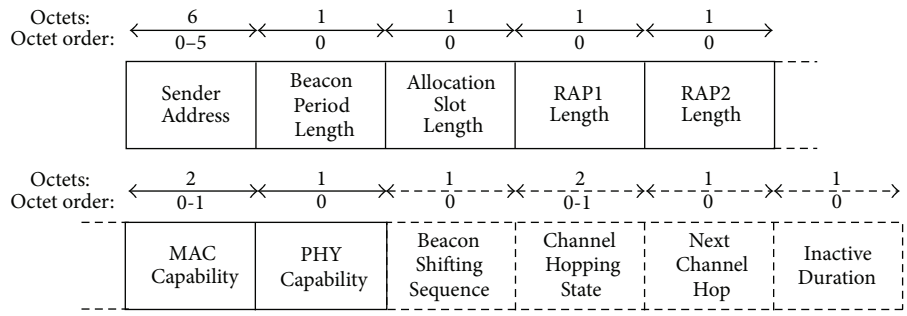


Figure 11: WBAN beacon frame format.

purpose, the WUSB/WBAN host has to allocate the WUSB private channels. Except for the RAP1 period, length of the other periods can be set to zero. By using this feature, the WBAN host which also performs the function of WUSB host allocates the WUSB private channels at the RAP2 period.

In the WUSB over WBAN architecture, in order to set up a wireless communication link to WHMS, secure WUSB Channels should be encapsulated within a WBAN superframe. This enables the MMC scheduling between WUSB host and its several peripheral devices without contention. Figure 16 shows the example topologies of the WUSB over WBAN architecture in both nonmedical and medical traffic environments. And Figure 17 explains A signaling flow for WUSB private channel reservation over WBAN protocol.

In this scenario, the user carries a portable or WHMS host device. This host device performs roles of the WUSB host and the WBAN hub simultaneously. Therefore, a "wearable" WUSB cluster and a WBAN cluster can be formed. The attached input sensor nodes perform the functions of localization-based input interfaces for WHMS's healthcare monitoring. Furthermore, the attached wireless nodes comprise the peripherals of a wearable computer system, and the central WUSB host exchanges data with the outer peripherals of the WUSB slave devices.

5. Time-Synchronization and Localization Middleware in the Proposed WHMS

Many protocols have been reported for clock synchronization in WSNs, and these protocols all have some common basic approaches. Synchronization is achieved by exchanging clock (timestamp) information among nodes while reducing the effect of nondeterministic factors in message delivery. They can be classified into three types: receiver-receiver (R-R) synchronization, sender-receiver (S-R) synchronization, and one-way message dissemination.

In R-R synchronization, a node periodically broadcasts wireless beacon messages to its neighbors. The receivers use the message's arrival time as a point of reference for comparing their clocks, and then they exchange the local timestamps at which they received the same broadcast message, as shown in Figure 18(a). The receivers finally compute their offset based on the difference in reception times to synchronize their clocks. Reference broadcast synchronization (RBS) [2] is a typical R-R synchronization protocol.

S-R synchronization is performed by a handshake protocol between a pair of nodes. Examples of synchronization protocols that employ this approach include timing-sync protocol for sensor networks (TPSN) and tiny-sync and minisync (TS/MS) [7]. Figure 18(b) shows a two-way message exchange mechanism in TPSN protocol. Node B sends a message at its local time T_1, and Node A receives this packet at its local time T_2. At time T_3, Node A sends back an acknowledgment packet which contains the values of T_2 and T_3. After receiving the packet at T_4, Node B can calculate the clock offset (θ) and propagation delay (d) as in (1). Knowing the offset, Node B can synchronize its clock to Node B's clock by adding θ to its current clock value.

This is expressed in the following equation:

$$\theta = \frac{(T_2 - T_1) - (T_4 - T_3)}{2}, \qquad d = \frac{(T_2 - T_1) + (T_4 - T_3)}{2}. \tag{1}$$

In one-way message dissemination, a reference node broadcasts its timing information to its neighbors, and they record the arrival times of the broadcast message, as shown in Figure 18(c). Collecting all the timestamps, each node can convert between the local hardware clock and the clock of the reference node by linear regression table. Flooding time-synchronization protocol (FTSP) [7] is a typical protocol utilizing one-way message dissemination scheme.

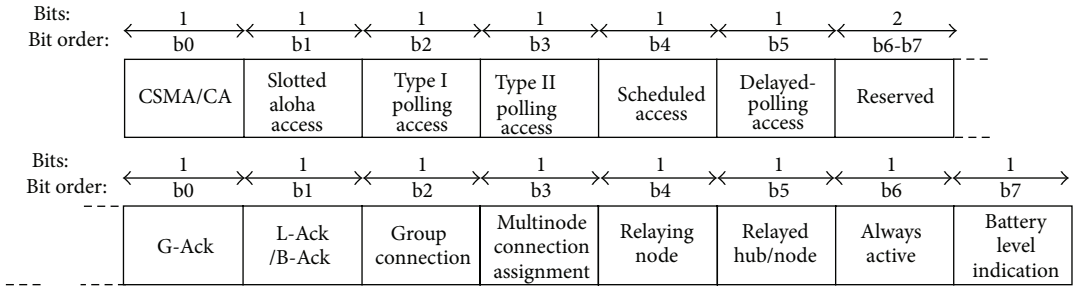


Figure 12: Current WBAN MAC Capability format standard.

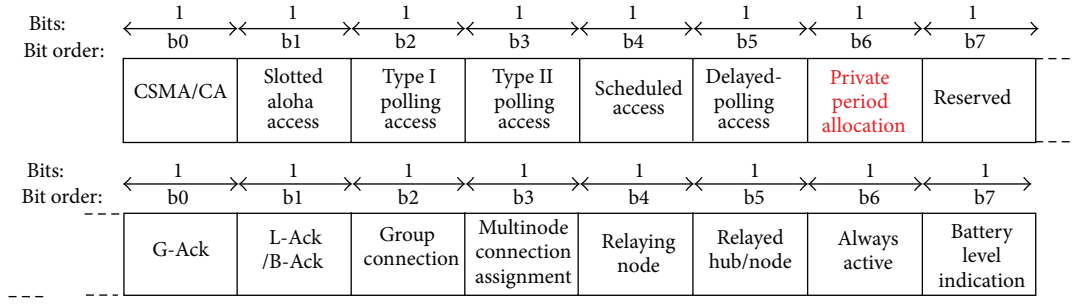


Figure 13: A new WBAN MAC Capability format for WUSB private channel allocation.

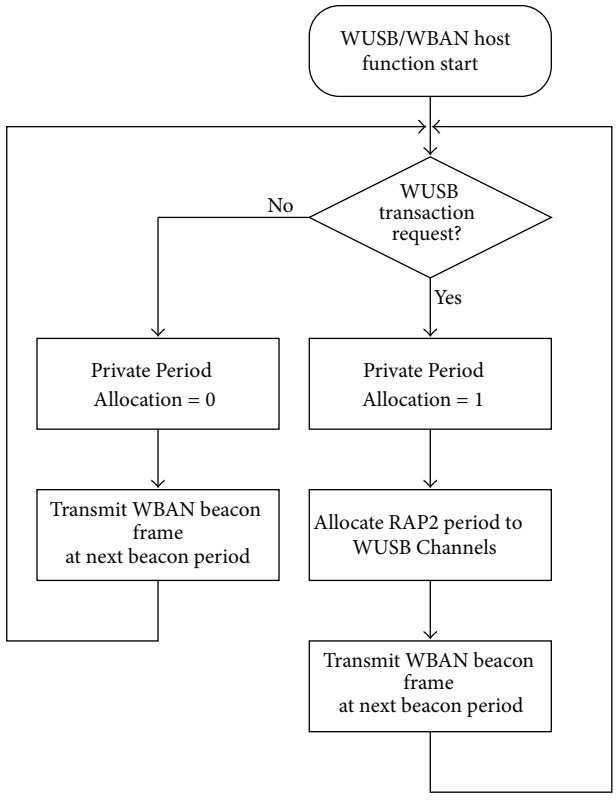


Figure 14: WBAN beacon signaling framework (BSF) for WUSB private channel allocation.

A clock is a device that measures time. It generally consists of a periodic component such as an oscillator and a counting component. The characteristics of the oscillator and the counter define the clock's behavior. The frequency of the oscillator controls the rate at which the clock advances. Since it is impossible to create oscillators that oscillate at exactly the same rate, every clock advances at a different rate in real world.

The time reported by a clock at some ideal time t is written as $C(t)$. We will write $C_A(t)$ as the time given by the local clock of Node A. The difference between the time of an ideal clock and a given clock is said to be the offset, $\theta(t)$, which is defined as in (2). The relative offset between Node A and Node B from the viewpoint of Node B, $\theta_B^A(t)$, is defined in (2).

The oscillator in a clock produces periodic pulses. The difference between the rate these pulses are produced at and the rate an ideal clock counts the desired interval is called the skew (frequency offset). Let $\varepsilon(t)$ denote the skew as in (2). The clock skew is the slope of the change in offset compared to the local clock. The slope of the relative offset $\theta_B^A(t)$ is relative skew $\varepsilon_B^A(t)$. This value means the skew of Node A's timer from the viewpoint of Node B's timer and is defined as in (2).

Typical quartz oscillators exhibit frequency offsets on the order of a few parts per million (ppm). For example, a 20 ppm oscillator will introduce an error of maximum 72 ms in 1 hour. For the remainder of the paper, we adopted the linear clock skew model so that all the notations of skew (ε) are used without the term of time (t). Although the skew may be time varying [7], modeling a frequency offset as a piecewise linear

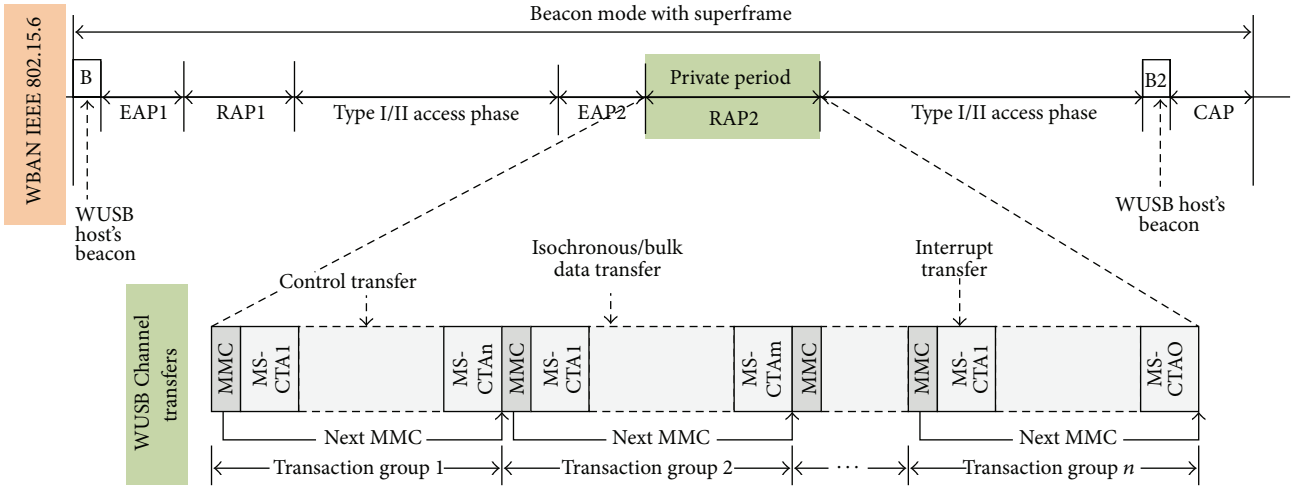
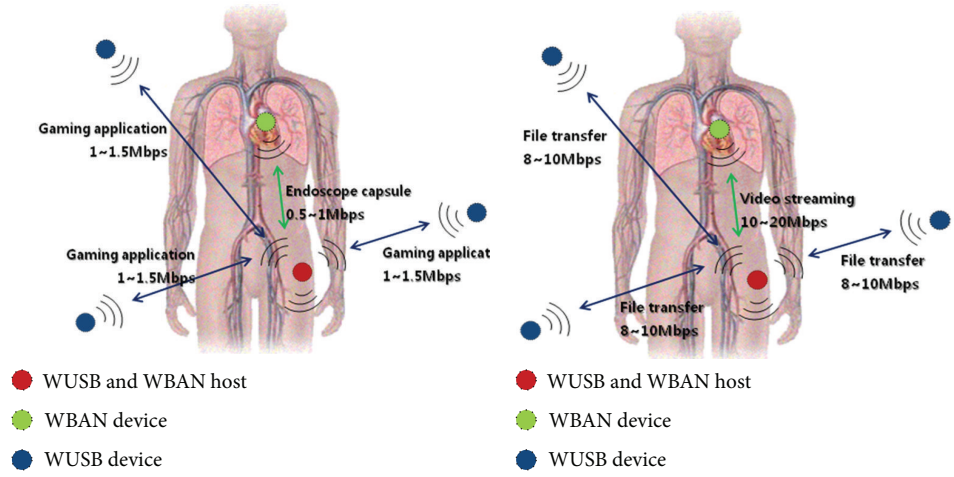


FIGURE 15: IEEE 802.15.6 superframe structure for WUSB private channel allocation.



(a) Medical service scenario (b) Nonmedical service scenario

FIGURE 16: WUSB over WBAN architecture.

function of time is a reasonable step as it changes relatively slow. If we refer the operating frequency of Node B's clock to f_B, f_A is expressed as in (2) and the relative skew ε_B^A have another form of definition as follows:

$$\theta(t) = t - C(t), \qquad \theta_B^A(t) = C_A(t) - C_B(t),$$

$$\varepsilon(t) = \frac{1}{C'(t)} - 1,$$

$$\varepsilon_B^A(t) = \frac{C_A'(t)}{C_B'(t)} - 1 = \frac{\varepsilon_B(t) + 1}{\varepsilon_A(t) + 1} - 1, \tag{2}$$

$$f_A = f_B\left(1 + \varepsilon_B^A\right), \qquad \varepsilon_B^A = \frac{f_A}{f_B} - 1.$$

In the context of WSNs, time synchronization refers to the problem of synchronizing clocks across a set of sensor nodes which are connected to one another over single-hop or multi-hop wireless networks. The time-synchronization problem in WSNs generally involves two steps. The first is synchronizing the nodes in the network to one common absolute time by adjusting the clock offset among the nodes, and the second is correcting the clock skew relative to a certain standard frequency. The latter is because the imperfections in the quartz crystal and the environmental conditions cause different clocks to run at slightly different frequencies. The effect of clock skew is the main reason that clock offset keeps drifting away. Adjusting clock skew guarantees long-term reliability of synchronization and it reduces networkwide energy consumption in synchronization protocols.

Figure 19 shows the synchronization process of Node B's clock to synchronize with Node A's clock. We assume that the two clocks have the identical initial value for the sake of simplicity. Since each oscillator has its unique clock frequency, the clock offset between the two nodes keeps increasing. If Node B determines the relative offsets $\theta_B^A(t)$ and executes offset compensation at t_1, t_2, t_3, and t_4, the corresponding

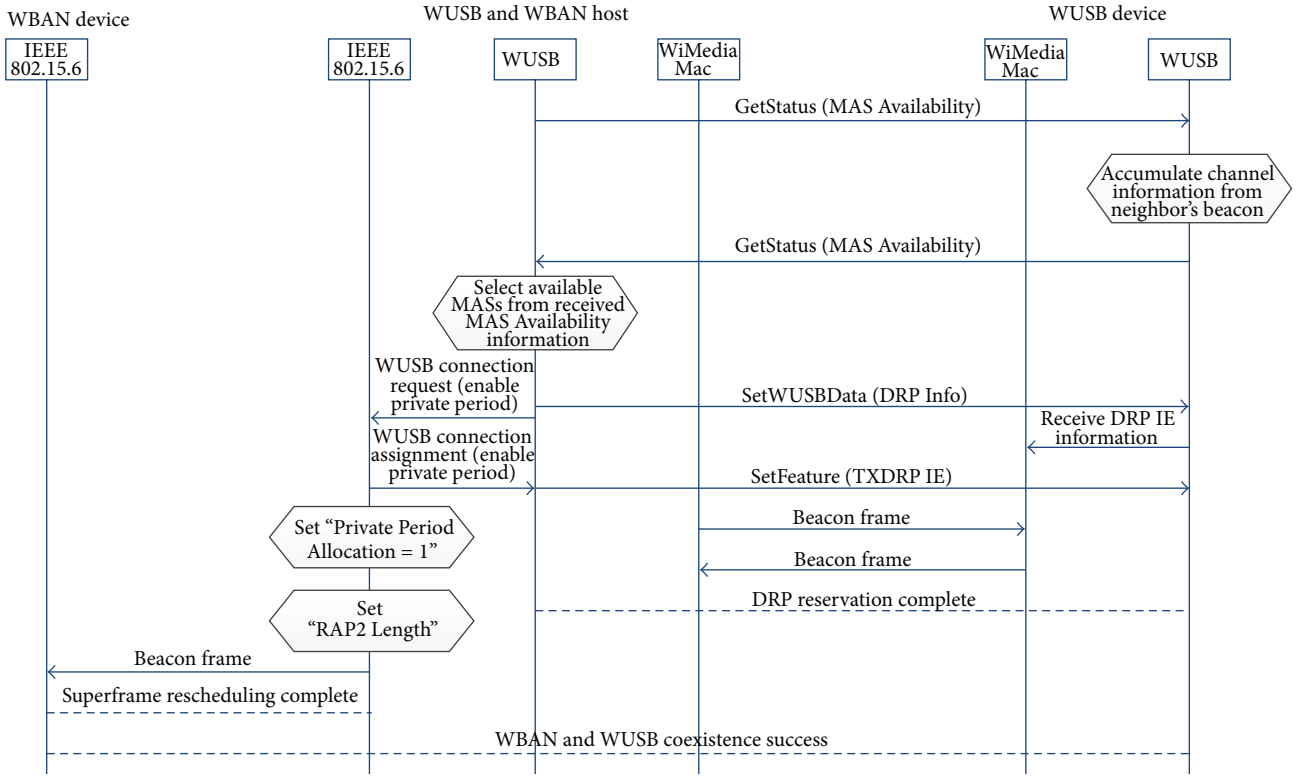
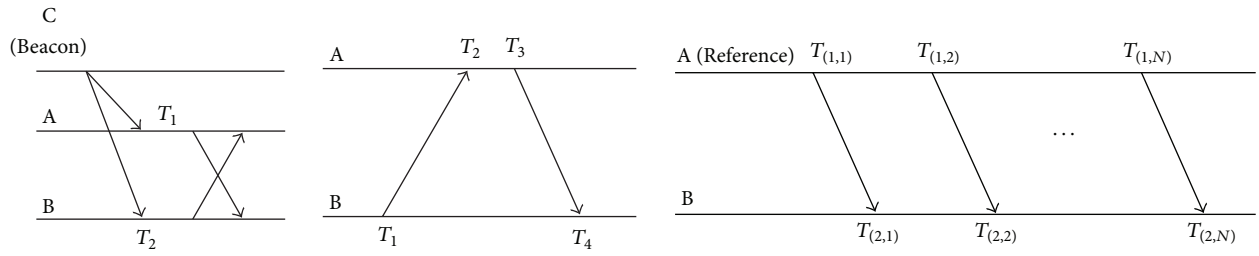


Figure 17: A signaling flow for WUSB private channel reservation over WBAN protocol.



(a) R-R synchronization (b) S-R synchronization (c) One-way message dissemination

Figure 18: Basic synchronization approaches.

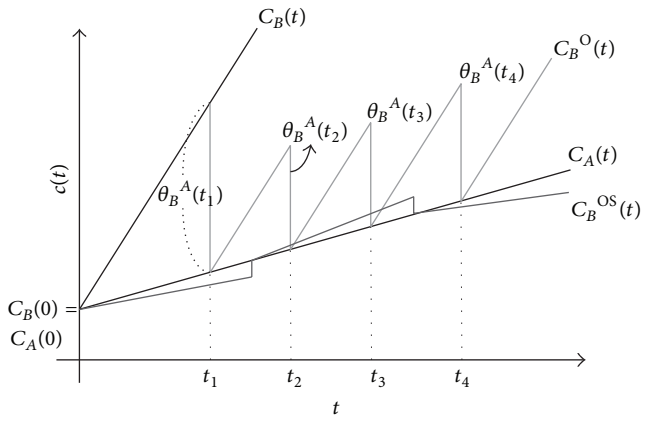


Figure 19: Offset compensation and skew compensation.

synchronized clock will be $C_B^O(t)$. Since the offset compensation does not take the impact of clock drift into account, $C_B^O(t)$ keeps the same varying rate and drifts away from $C_A(t)$. If the synchronization (offset compensation) interval becomes longer, the synchronization error also becomes larger. The solution for improving synchronization accuracy is to decrease the synchronization period, but frequent resynchronization will introduce too much communication overhead. If the relative clock skew $\varepsilon_B^A(t)$ is accurately estimated, the corresponding synchronized clock can be $C_B^{OS}(t)$ as in Figure 19. $C_B^{OS}(t)$ approximately approaches $C_A(t)$ although it does not completely synchronize with $C_A(t)$. However, it is difficult to estimate the skew accurately due to various factors including short-term effects such as supply voltage, temperature, and humidity changes and long-term

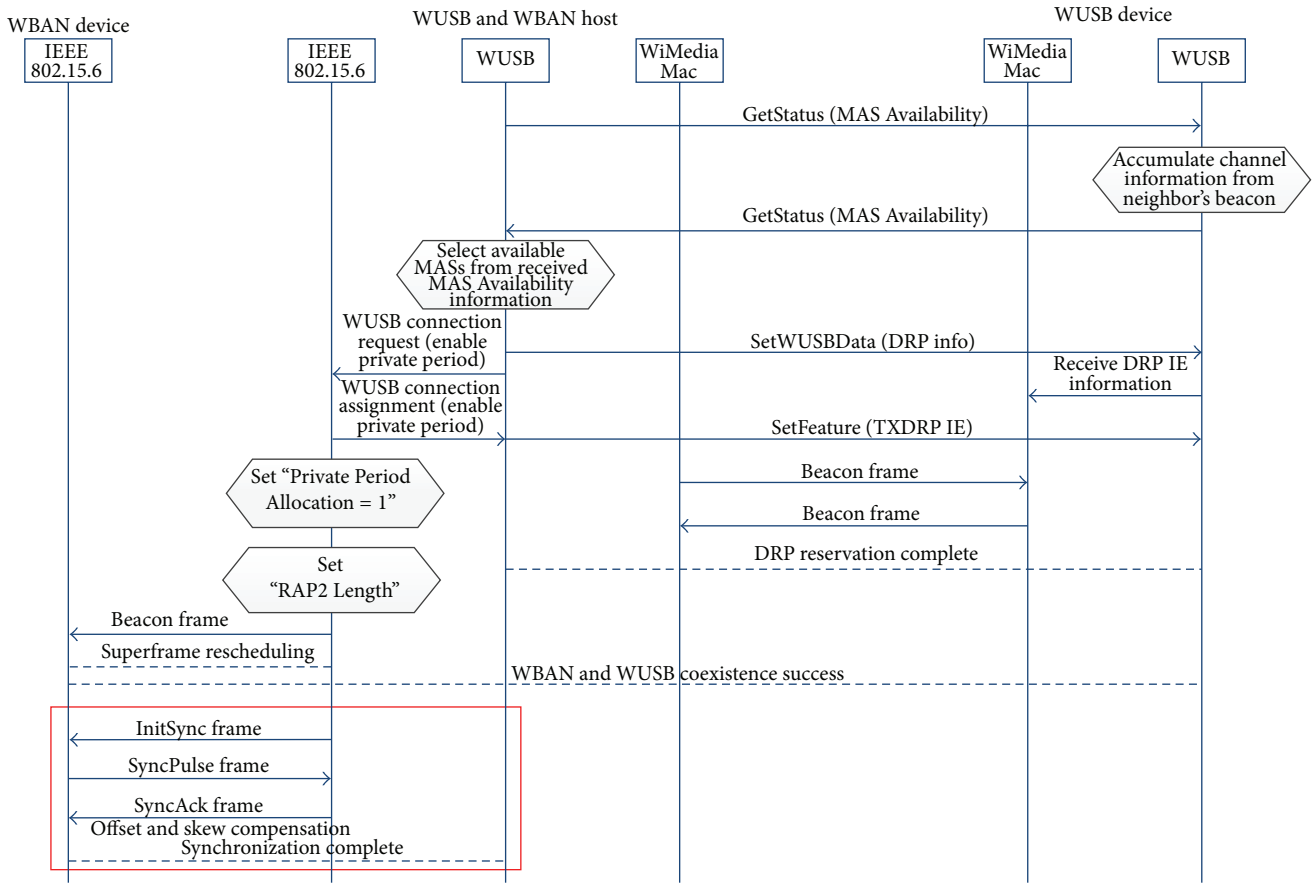


Figure 20: Procedures for time synchronization in WUSB over WBAN protocol.

effects such as crystal aging. Offset compensation enables short-term time synchronization, while skew compensation facilitates long-term synchronization. Therefore, these two compensation techniques should be employed together, so that the time-synchronization algorithm with high accuracy and moderate communication overhead can be realized. By using this time-synchronization technique, procedures for time-synchronization in WUSB over WBAN protocol are proposed in Figure 20.

We have built a WUSB over WBAN test bed for the WHMS workspace shown in Figure 21. The WBAN host uses the standard time difference of arrival technique by observing the time lag between the arrival of the signals and estimates its distance from each WUSB/WBAN biosensor device. The estimated distances are passed to the context-aware WHMS server, which computes the location of the WBAN biosensor device using the distances. WHMS application service is requested by providing the WUSB/WBAN host with the user's location information.

The context-aware WHMS application server computes the location (x, y, z) of the WBAN biosensor device using the distances through the trilateration method, as illustrated in Figure 22. Suppose the location of the ith WBAN biosensor device is denoted by b_i, and the estimated distance between the ith WBAN biosensor device and the WUSB/WBAN host is d_i. Then, we made the following equation for the ith WBAN biosensor device as in (3). This equation is solved using Newton-Raphson method [8]. The initial location (x_0, y_0, z_0) of the WUSB/WBAN host is set to the center of the workspace. Now consider the following:

$$(x - b_{ix})^2 + \left(y - b_{iy}\right)^2 + (z - b_{iz})^2 = {d_i}^2. \quad (3)$$

6. Results and Discussion

Performance of the proposed WHMS scheme is evaluated through NS2 simulations with WBAN PHY/MAC simulation parameters [9–11]. The network size is 5 m $*$ 5 m. Maximum 20 devices are randomly deployed into this workspace area. WBAN frame size is fixed to 4095 bytes. Figure 23 shows throughputs of a non-WUSB WBAN device according to each UWB PHY data rate for each m_{In} situation. The m_{In} parameter indicates the occupation ratio of WUSB traffic for the entire WBAN superframe length. In Figure 23, the larger m_{In} ratio of WUSB traffic reduces throughput of a non-WUSB WBAN device more. This result is caused by the reason that the increase of m_{In} ratio enlarged length of the RAP2 period allocated for WUSB traffic in the WBAN superframe.

In Figure 24, UWB PHY data rate of devices is fixed to 480 Mbps. As shown in Figure 24, throughput of a non-WUSB WBAN device does not vary largely according to the frame size after the frame size exceeds a certain threshold.

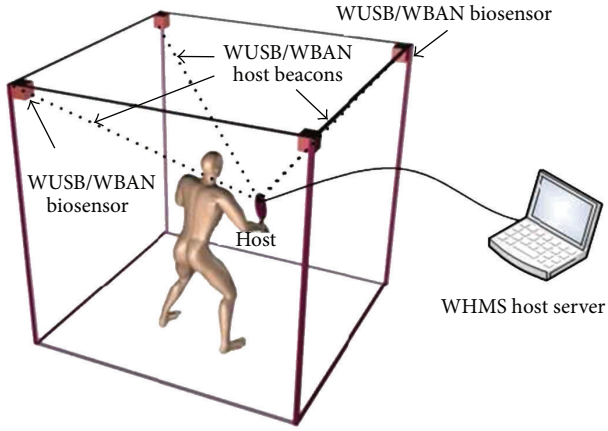


Figure 21: Test bed of a WHMS workspace.

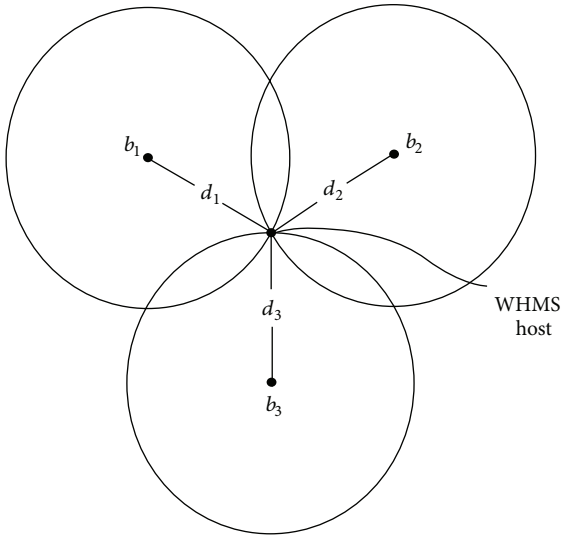


Figure 22: Trilateration method.

Because the WUSB transactions in RAP2 period are performed privately without contention, the increase of frame size does not cause the increase of frame collision probability. But the throughput of a non-WUSB WBAN device decreases extremely and proportionally according to m_{In} ratio of WUSB traffic.

To validate our time-synchronization middleware, the WBAN biosensor nodes use the 32.768 kHz real-time clock, and the frequency tolerance is set to ±20 ppm. We performed measurements on the WHMS test bed. A WBAN biosensor node acts as the reference node. Another WUSB/WBAN host node acts as the sink node which is connected to WHMS server and gathers information about the other WBAN biosensor nodes' synchronization error values. We measured the synchronization errors between the reference WBAN biosensornode and the other WBAN biosensor nodes.

We built a network which consists of 20 WBAN biosensor nodes. To measure the synchronization precision, we make each sensor node report the clock offset (θ_i) to the WUSB/WBAN host sink node whenever it performs offset

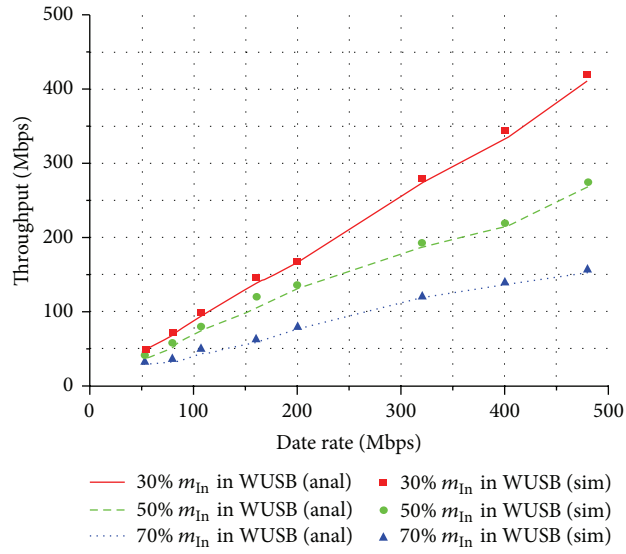


Figure 23: Throughput of a non-WUSB WBAN device according to each UWB PHY data rate.

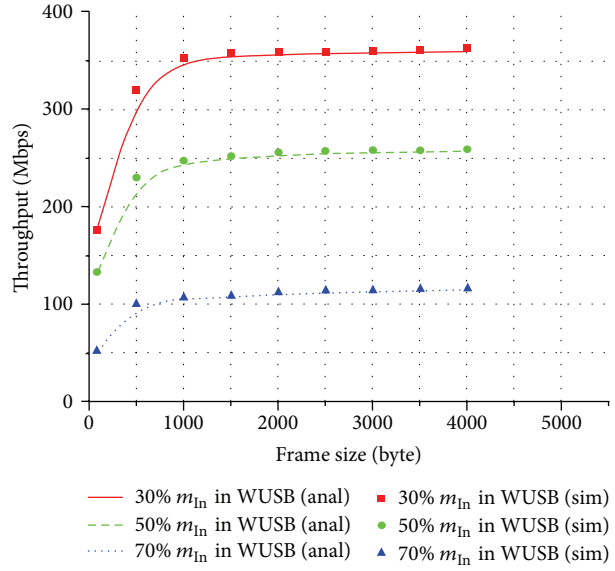


Figure 24: Throughput of a non-WUSB WBAN device according to each frame size.

compensation. This clock offset expresses the time difference between its own clock and the clock of reference node at the moment just before offset compensation. However, this value does not coincide with the synchronization error exactly since it may introduce some measurement errors. For the offset compensation, we adopt the S-R synchronization scheme of two-way message exchanges in TPSN. The WUSB/WBAN host sink node controls the initiation of two-way message exchanges and its interval. Once initialized by the WUSB/WBAN host sink node, the synchronization

Table 2: Estimated distances at 100 cm.

Distance (cm)	Frequency (2000 times in total)
99.9674	11
100.001	327
100.035	946
100.069	62
100.104	573
100.138	75
99.071	6

Table 3: Statistical analysis at 100 cm.

Statistics	Value
Mean	100.041 cm
Variance	0.00154603 cm
Standard deviation	0.03931958 cm

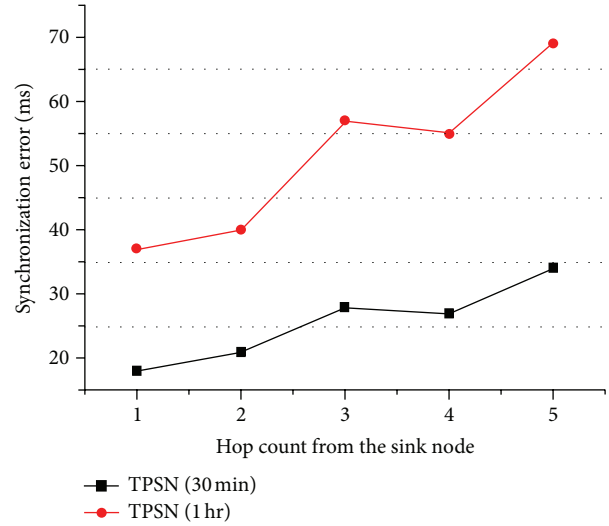


Figure 25: Synchronization error by offset and skew compensation.

procedure of exchanging two-way message can be driven to the whole network.

For checking the effect of clock skew, we allowed WBAN biosensor nodes to perform only offset compensation as in the original TPSN. Every WBAN biosensor node performed offset compensation at the interval of 30 minutes and 1 hour. The data reported from each WBAN biosensor node are averaged according to message hop count. Figure 25 shows this result of the average synchronization error. It is checked that the synchronization errors weakly tend to increase with respect to message hop count from the WUSB/WBAN host sink node and that the effect of clock skew on synchronization error becomes considerable as time passes.

Finally, to validate our localization middleware, the distances between the WBAN biosensor nodes and the WUSB/WBAN host node are obtained in an asynchronous mode. All WBAN biosensor nodes use a 16-bit timer, and the location of a WBAN biosensor node is computed at 10 Hz. The accuracy of the distance estimation has been tested. For a fixed position of the WUSB/WBAN host, the distance from a WBAN biosensor node is measured by hand. Then, the distance is estimated in the proposed middleware framework. Such an estimation is done for 2,000 times. Tables 2 and 3 show the statistics for the measured distances of 100 cm. This distance test proves that our localization middleware can be successfully integrated with WHMS applications.

7. Conclusion

In this paper, we propose time-synchronization and localization middlewares platform built on WUSB over IEEE 802.15.6 WBAN hierarchical protocol. Our test results prove that they can be well integrated and lead to a new type of natural interface for the wearable health-monitoring systems (WHMS). In the current implementation, the location data are computed at 10 Hz. For fully supporting real-time applications of WHMS, more effort should be made to increase the performance. Further, many applications may benefit not only from location awareness but also from orientation awareness built on the precise time synchronization. To fulfill such needs, the overall performance should be continuously upgraded.

Acknowledgments

This research was supported in part by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (MEST) (2010-0002366) and in part by Midcareer Researcher Program through NRF Grant funded by the MEST (2011-0016145).

References

[1] A. Pantelopoulos and N. G. Bourbakis, "A survey on wearable sensor-based systems for health monitoring and prognosis," *IEEE Transactions on Systems, Man and Cybernetics C*, vol. 40, no. 1, pp. 1–12, 2010.

[2] M. Patel and J. Wang, "Applications, challenges, and prospective in emerging body area networking technologies," *IEEE Wireless Communications*, vol. 17, no. 1, pp. 80–88, 2010.

[3] IEEE 802.15 WPAN Task Group 6 Body Area Networks (BAN). IEEE, http://www.ieee802.org, 2010.

[4] Certified Wireless USB 1.1. USB-IF, http://www.usb.org, 2010.

[5] J. W. Kim, K. Hur, J. Park, and D. S. Eom, "A distributed MAC design for data collision-free wireless USB home networks," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 3, pp. 1337–1343, 2009.

[6] R. Palit, A. Singh, and K. Naik, "An architecture for enhancing capability and energy efficiency of wireless handheld devices," *International Journal of Energy, Information and Communications*, vol. 2, pp. 117–136, 2011.

[7] H. Dai and R. Han, "TSync: a lighweight bidirectional time synchronization service for wireless sensor networks," *ACM*

SIGMOBILE Mobile Computing and Communications Review, vol. 8, pp. 125–139, 2004.

[8] N. B. Priyantha, A. Chakraborty, and H. Balakrishnan, "Cricket location-support system," in *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking (MOBICOM '00)*, pp. 32–43, August 2000.

[9] N. Karthikeyan, V. Palanisamy, and K. Duraiswamy, "Performance comparison of broadcasting methods in Mobile Ad Hoc Network," *International Journal of Future Generation Communication and Networking*, vol. 2, pp. 47–58, 2009.

[10] K.-I. Kim, "Adjusting transmission power for real-time communications in wireless sensor networks," *Journal of Information and Communication Convergence Engineering*, vol. 10, pp. 21–26, 2012.

[11] M. Mana, M. Feham, and B. A. Bensaber, "SEKEBAN (secure and efficient key exchange for wireless body area network)," *International Journal of Advanced Science and Technology*, vol. 12, pp. 45–60, 2009.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 521383, 7 pages
http://dx.doi.org/10.1155/2013/521383

Research Article

Robust People Tracking Using an Adaptive Sensor Fusion between a Laser Scanner and Video Camera

Yeong Nam Chae,[1] Yeong-Jae Choi,[1] Yong-Ho Seo,[2] and Hyun S. Yang[1]

[1] *Department of Computer Science, KAIST, 291 Daehak-ro, Yuseong-gu, Daejeon 305-701, Republic of Korea*
[2] *Department of Intelligent Robot Engineering, Mokwon University, 88 Doanbok-ro, Seo-gu, Daejeon, Republic of Korea*

Correspondence should be addressed to Yong-Ho Seo; yhseo@mokwon.ac.kr

Received 25 January 2013; Accepted 28 February 2013

Academic Editor: Sabah Mohammed

Robust detection and tracking in a smart environment have numerous valuable applications. In this paper, an adaptive sensor fusion method which automatically compensates for bias between a laser scanner and video camera is proposed for tracking multiple people. The proposed system comprises five components: blob extraction, object tracking, scan data clustering, a cluster selection, and updating the bias. Based on the position of object in an image, the proposed system determines the candidate scan region. Then, the laser scan data in the candidate region of an object is clustered into several clusters. A cluster which has maximum probability as an object is selected using a discriminant function. Finally, a horizontal bias between the laser scanner and video camera is updated based on the selected cluster information. To evaluate the performance of the proposed system, we show error analysis and two applications. The results confirm that the proposed system can be used for a real-time tracking system and interactive virtual environment.

1. Introduction

Robust detection and tracking in a smart environment have numerous valuable applications, such as recognizing human behavior for intelligent surveillance, monitoring, and analyzing. To monitor human activities such as location, identity, and behavior, robust detection and tracking are necessities in various environments.

In the people motion tracking area, the one side view, which comes from a single video image, is used to verify peoples' actions. In diverse and sophisticated environments, there are numerous problems in a single video image. In ubiquitous environments, to perceive the motion of people, unfavorable conditions exist in single-sensor systems, such as illumination variation and shadow. To overcome these problems, many methods rely on fusing a number of sensors, such as the infrared cameras, laser range finders, and image cameras of many directions [1, 2]. However, as various sensors are added, the object calibration is a very important issue in the object tracking area for the reliable detection and tracking of multiple objects [3].

To calibrate multi-sensors, an attempt using an extrinsic calibration between the camera and the laser scanner was proposed in [3–6]. This approach uses a technique in which both sensors image a planar checkerboard target at unknown orientations. Even if the calibration is reliable, it is inconvenient to adjust for tracking people simultaneously. Furthermore, the checkerboard always should be used to calibrate this system.

To overcome these problems, we present an adaptive sensor fusion method between the laser scanner and video camera. In this proposed approach, our system does not need a checkerboard. As the configuration of the system, which consists of a laser scanner and video camera, changes to intuitive positions, the traditional system requires a recalibration process. The proposed system, which has a background model, can compensate these sensor's horizontal variations automatically.

The paper is organized as follows: in the next section, we briefly introduce the proposed system. In Section 3, we present a video processing method for tracking multiple

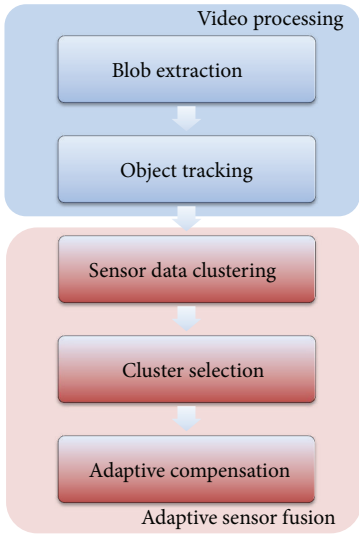


Figure 1: System overview.

objects. In Section 4, we present a detailed method to calibrate laser scanner and video camera. Then, we present the results with real environmental data in Section 5. At last, we conclude in Section 6.

2. System Overview

The proposed system consists of the video processing and the adaptive sensor fusion as shown in Figure 1. To measure the position of objects in an image, we adopt Mixture of Gaussians-(MoGs-) [7] based blob extraction and inference-graph-based object tracking approach. Then laser, scan data in the candidate region from the result of the object tracking is clustered into several clusters. A cluster which has maximum probability as an object is selected using a discriminant function. Finally, a horizontal bias between the laser scanner and video camera is compensated based on selected cluster information.

3. Video Processing

In order to merge the laser scanner and video camera, the proposed system performs video processing in a prior part of the system. In order to measure the position of objects in an image, we developed an enhanced view-based multiple-object tracking system based on previous research [8, 9]. In this section, we briefly introduce this multiple-object tracking system.

3.1. Blob Extraction. For segmentation, MoGs is widely used due to the capability for adaptation to the various environmental changes like illuminations. But, the main problem of MoGs is that moving objects are learned as background when they stop. And thus it fails to segment objects stopped as foreground. In the proposed system, in order to segment an input image into a foreground and background, a modified MoGs [8] is used to keep moving objects segmented as foreground, even when they stop. After conventional MoGs is performed, the following four steps are further executed to manage objects stopped. First, the Gaussians are sorted in descending order of their weight. Second, the pixels moving from the first Gaussian component to the second Gaussian component are identified as pixels belonging to the objects stopped and they are put into an augmented mask. Third, the pixels in the augmented mask are added to the segmentation mask from the conventional MoGs. Fourth, the pixels in the augmented mask are removed from the augmented mask when the stopped objects start to move. This modified MoGs guarantees that the objects still identified as foreground, even when moving objects stopped. In order to remove shadows and highlights, we can adjust the intensity of a shadow pixel compared with a background model as shown in Figure 2(d). After removing shadows and highlights, when capturing foreground pixels at frame t, they are clustered into a set of $B^t = \{b_i^t \mid i \text{ is an integer and } 0 \le i\}$, where a blob b_i^t is an ith set of connected foreground pixels.

3.2. Object Tracking. A blob represents an object in the ideal case. In the real environment, however, one object can have several blobs (fragmentation), and one blob can have multiple-objects (grouping). To deal with these problems, Choi et al. [9] adopt an online multiple object tracking framework. Figure 3 shows the overall procedure of this framework. First, detecting blob association events between B^t and B^{t-1} can update the blob inference graph, labeling each vertex as fragment, object, and group. Finally, localization of objects can be captured by using the blob graph.

4. Adaptive Sensor Fusion

In this paper, we propose an adaptive method which compensates horizontal bias between the laser scanner and video camera. In order to merge the laser scanner and video camera adaptively, there are three steps in the proposed method. First, laser scan data in the candidate region of the object is clustered into several clusters. Second, a cluster which has maximum probability as an object is selected using a discriminant function. Finally, a horizontal bias between the laser scanner and video camera is updated based on selected cluster information.

4.1. Sensor Data Clustering. To match an object in an image to laser scan data, we determine the candidate region in the laser scan data. The candidate region in the laser scan data can be determined based on the object position in an image. This candidate region (θ_l, θ_r) is shown in the following equations:

$$\theta_l = (1 - \varphi)\left(90 + \text{fov}\left(-\frac{1}{2} + \frac{o_l}{\omega}\right) + \text{bias}\right),$$
$$\theta_r = (1 + \varphi)\left(90 + \text{fov}\left(-\frac{1}{2} + \frac{o_r}{\omega}\right) + \text{bias}\right). \quad (1)$$

In (1), θ_l and θ_r are the left and right angle boundaries, respectively, o_l denotes the left pixel position of an object, o_r denotes the right pixel position of an object, ω refers to image

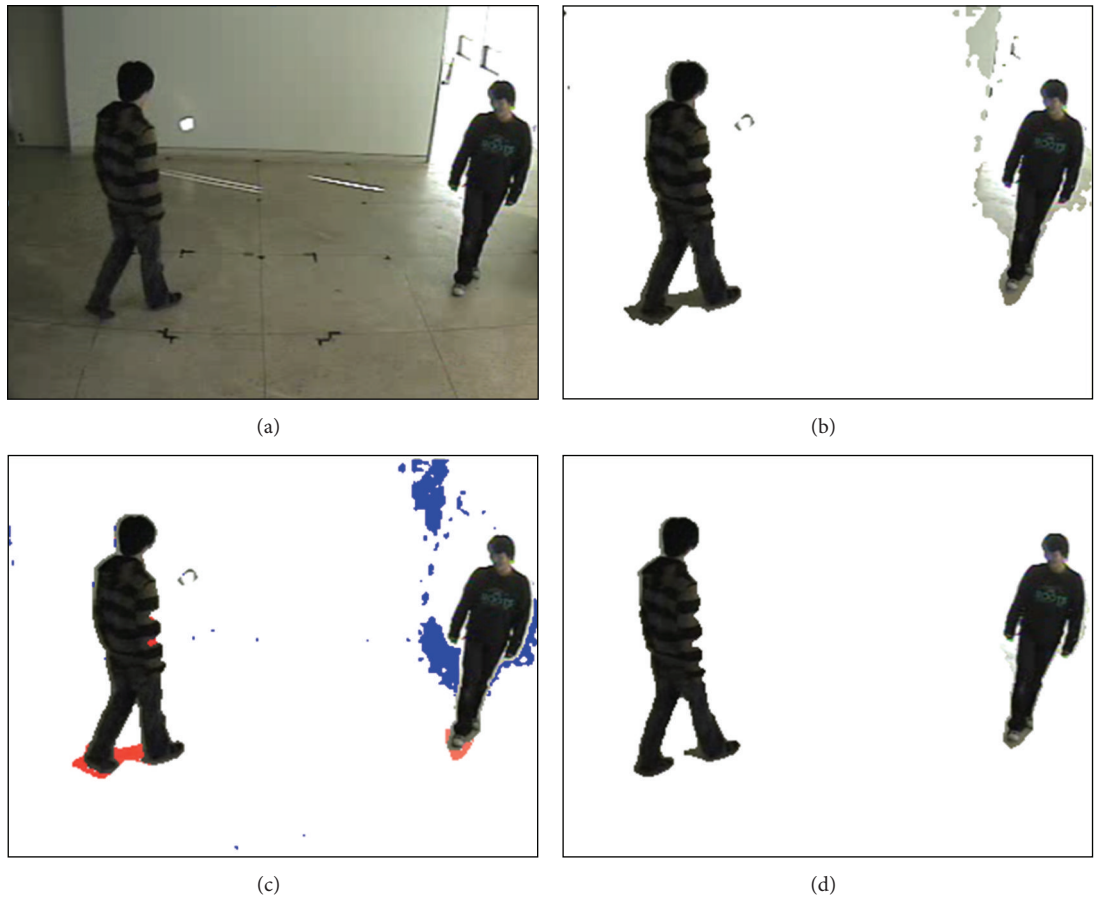


FIGURE 2: (a) An input image. (b) The result of foreground extraction. (c) Detected shadow (marked in red) and highlight (marked in blue). (d) The result of foreground extraction: an example of a figure without shadow or highlight.

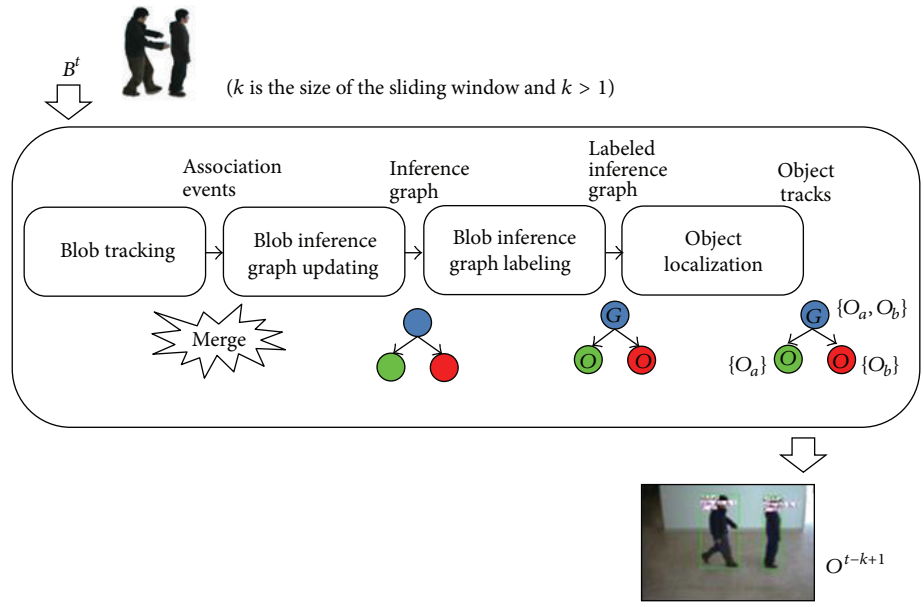


FIGURE 3: The online multiple-object tracking framework.

width, bias means horizontal bias between the laser scanner and video camera, fov means field of view angle, which is determined by the video camera, and φ denotes the scale factor which expands the candidate region.

The archived laser scan data in the candidate region is lumped into several clusters by the nearest neighborhood clustering algorithm. We use the following equation as a threshold equation:

$$\sqrt{d_{\mathrm{MV}}(\theta_1)^2 + d_{\mathrm{MV}}(\theta_2)^2 + 2d_{\mathrm{MV}}(\theta_1)\, d_{\mathrm{MV}}(\theta_2) \cos(|\theta_1 - \theta_2|)} > \delta. \tag{2}$$

In (2), $d_{\mathrm{MV}}(\theta)$ refers to the distance from the measured value of the laser scanner at angle θ, both θ_1 and θ_2 are discrete angle values $(0, 1, 2, \ldots 180)$, and δ denotes the neighborhood distance threshold value, which is set as 2.5 feet (an average person's stride length [10]). Using this threshold function, we classified the datum in 2.5 feet from each other as same cluster.

4.2. The Cluster Selection Method. In order to take an appropriate cluster which has the greatest probability to be an object, the proposed system adapts the following discriminant function:

$$m = \arg\max_{0 \le i \le k} \sum_{\theta \in \mathbf{C}_i} \Delta(\theta). \tag{3}$$

In (3), k denotes the number of clusters, and C_i refers to the ith cluster. A discriminant function Δ is adapted to compare each cluster:

$$\Delta(\theta) = \begin{cases} \dfrac{\theta_\omega}{2} |\theta - \theta_c|\, d_{\mathrm{MAX\,BG}}, & d_{\mathrm{BG\,MV}}(\theta) > \delta, \\ \dfrac{\theta_\omega}{2} |\theta - \theta_c|\, d_{\mathrm{BG\,MV}}(\theta), & \text{otherwise.} \end{cases} \tag{4}$$

In (4), $d_{\mathrm{MAX\,BG}}$ denotes the maximum distance from the background model of the laser scan data, $d_{\mathrm{BG\,MV}}(\theta)$ denotes the difference between the background model and measured value at θ, θ_ω refers to angular width, and θ_c refers to the center angle of the candidate region (θ_l, θ_r). These are shown in the following equations:

$$d_{\mathrm{MAX\,BG}} = \max_{0 \le \theta \le 180} d_{\mathrm{BG}}(\theta), \tag{5}$$

$$d_{\mathrm{BG\,MV}}(\theta) = |d_{\mathrm{BG}}(\theta) - d_{\mathrm{MV}}(\theta)|, \tag{6}$$

$$\theta_\omega = |\theta_l - \theta_r|, \tag{7}$$

$$\theta_c = \frac{\theta_l + \theta_r}{2}. \tag{8}$$

In (5) and (6), $d_{\mathrm{BG}}(\theta)$ denotes the distance from the background model of the laser scan data at θ. The background model of the laser scan data is similar to the traditional background model of the gray scale video stream. Consider

$$D_{\mathrm{BG}}(\theta) = \begin{cases} d_{\mathrm{BG}}(\theta) + \alpha(d_{\mathrm{MV}}(\theta) - d_{\mathrm{BG}}(\theta)), & d_{\mathrm{BG\,MV}}(\theta) \ge \delta, \\ d_{\mathrm{BG}}(\theta) + \beta(d_{\mathrm{MV}}(\theta) - d_{\mathrm{BG}}(\theta)), & \text{otherwise.} \end{cases} \tag{9}$$

In (9), α refers to the learning rate in the case of foreground and β refers to the learning rate in the case of the background. Equation (9) is adapted to the whole range of laser scan data.

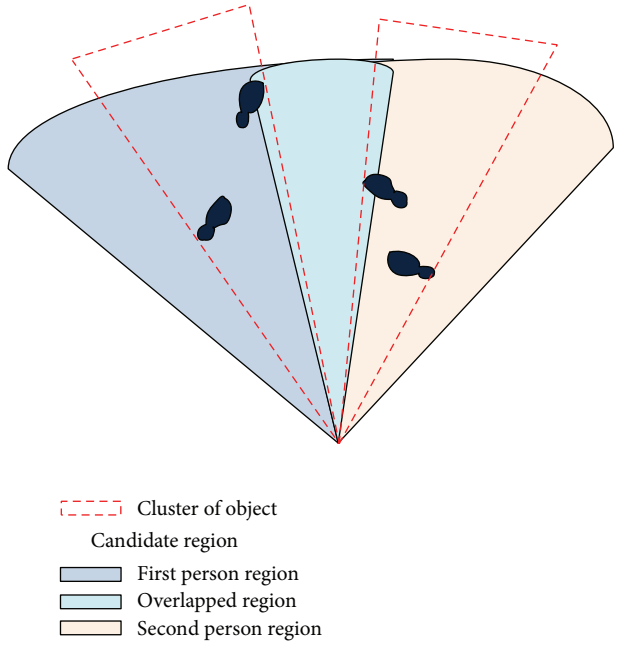


Figure 4: Example case of overlapping candidate regions.

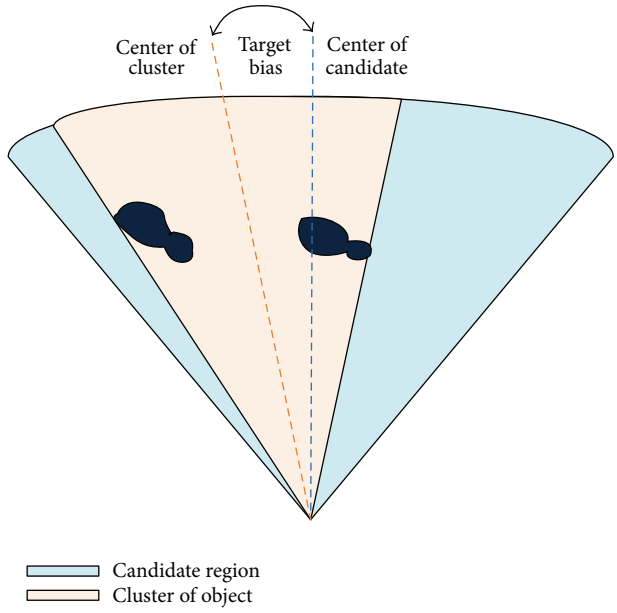
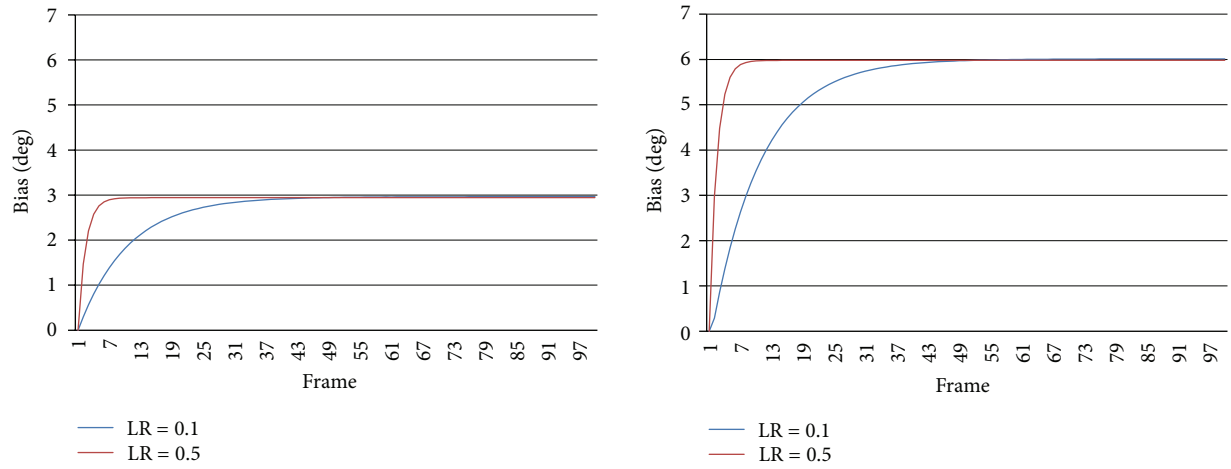


Figure 5: The horizontal bias angle.



(a) Degree 3 (b) Degree 6

Figure 6: The example of bias.

According to the discriminant function, the cluster volume, difference from the background model, and difference from the center of candidate region are considered as selection criteria. Using these selection criteria, the proposed system can select the appropriate cluster in case the candidate regions overlap. An example case of overlapping is shown in Figure 4.

4.3. Updating Rule. After determining an appropriate cluster, which is selected by a discriminant function, the proposed system updates bias adaptively. Figure 5 shows the target bias which is calculated from the candidate region and selected cluster. In order to take an appropriate cluster which has the greatest probability to be an object, the proposed system adapts the following discriminant function.

To diminish the target bias, the proposed system updates the bias using the following updating rule:

$$\text{bias} = \text{bias} + \rho\left(\text{MidAng}\left(m\right) - \theta_c\right). \tag{10}$$

In (10), bias refers to the horizontal bias angle between the field of view angle and the matched cluster, ρ denotes the learning rate, and MidAng(m) means the median angle of matched cluster m. The MidAng is calculated by the following equation:

$$\text{MidAng}\left(i\right) = \frac{\left|\min_{\theta\in C_i}\theta + \max_{\theta\in C_i}\theta\right|}{2}. \tag{11}$$

Using (10) as an updating rule, the proposed approach can compensate for horizontal bias adaptively.

5. Experimental Results

In order to evaluate proposed system, we arrange the video camera and the laser scanner vertically. Samsung SDC-415A model is adopted as video camera. It supports 768 × 494 resolution and covers 140 degree fields of view. SICK LMS100 model is adopted as laser scanner. It supports a 50 Hz scan rate over 270 degree range and 0.25 degree angular resolution. Its sensing range is 18 meters with an error of about 20 mm. The experimental results consist of two parts: the error analysis of the proposed approach and the application of the proposed system.

Table 1: The relative bias after convergence.

Learning rate	Angle			
	6	3	−3	−6
0.1	6.013	2.951	−3.139	−6.178
0.5	5.981	2.986	−3.080	−6.186

5.1. Error Analysis. To evaluate the proposed approach, we performed an error analysis of the bias compensation. To measure the error of the compensation, we installed a marker at the same position which can be detected by both the vision and laser scanner. By rotating the video camera three degrees horizontally, we measured the relative bias from the origin angle after convergence. The overall results are shown in Table 1. In this experiment, measurements were taken five times and the results were averaged out.

As shown in Table 1, the average error of the proposed approach is about 0.085 degree. The saturation of bias as time is shown in Figure 6. In Figure 6, LR refers to learning rate. In the case of 0.1, the saturation time is about 40 frames. However, in the case of 0.5, the saturation time is lower than 10 frames. From these results, the proposed approach can be considered reasonable for real-time systems.

5.2. Application. In this subsection, we show two applications of the proposed system. The first application is people tracking. The second application is a virtual pet system using augmented reality.

Figure 7 shows an example of multiple people tracking. The upper two images illustrate that the candidate regions do

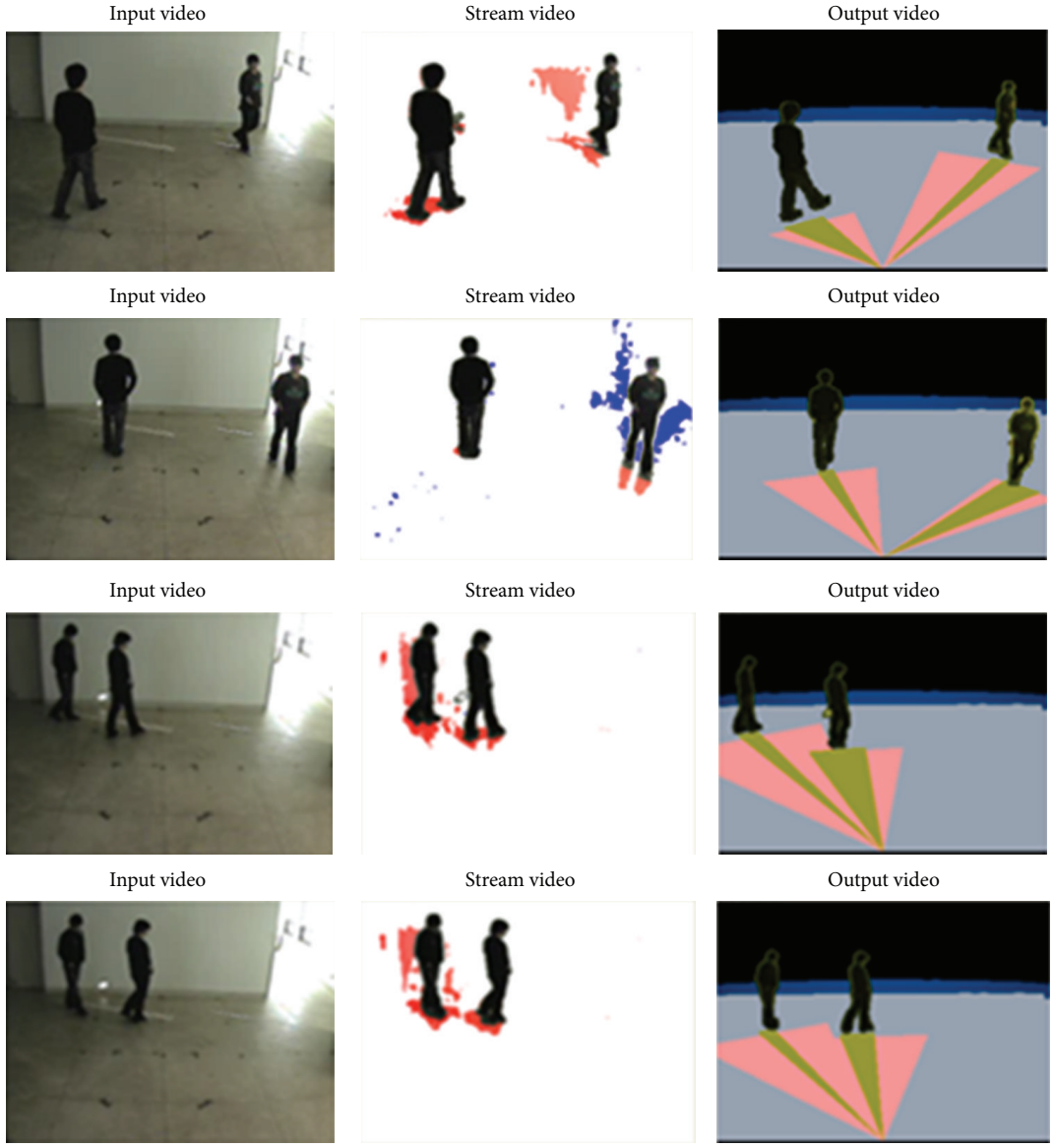


Figure 7: The people tracking by adjusting adaptive calibration.

not overlap, and the lower two images illustrate overlapping case. In both conditions, the proposed system can track appropriate people by the discriminant function.

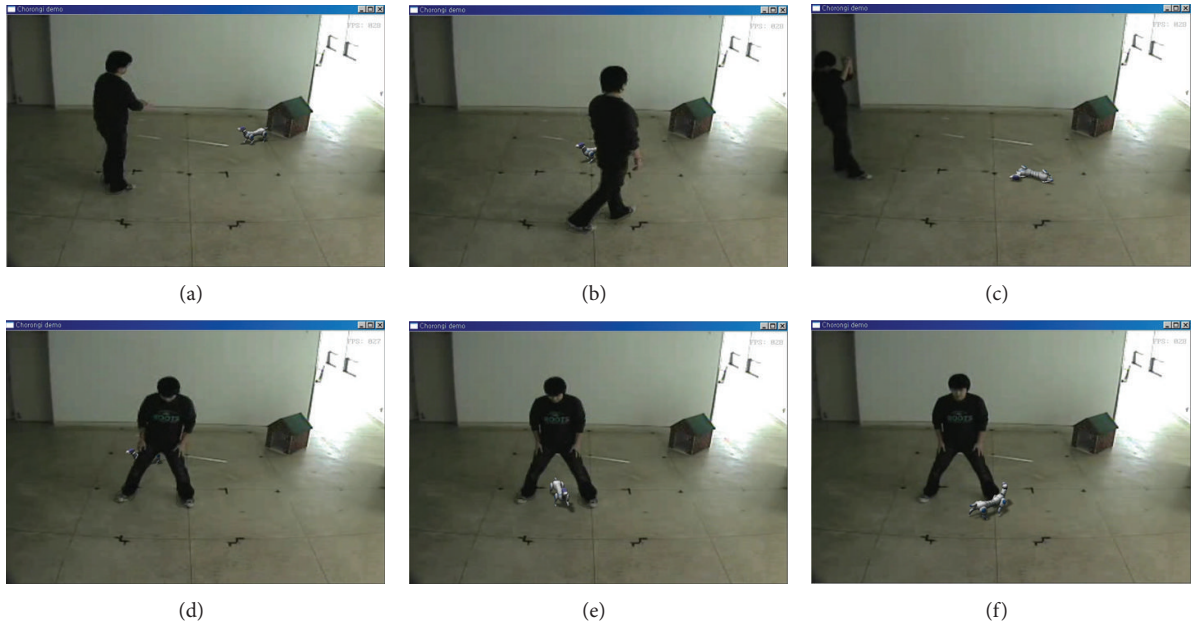
In the second application, to apply the proposed system, we employed it in a virtual pet system based on augmented reality. As shown in Figure 8, only a person and a dog's house exist in a real environment. With the position of the calibrated object result and the human action recognition result, user can make a relationship with a virtual pet, called Cho-Rong-I.

The ground of the real environment is calibrated in the initial stage so that the bottom-center coordinate of each object can be converted into the coordinate in the ground.

We made several complex scenarios in order to probe the performance of the proposed system as follows: Cho-Rong-I follows the owner in Figures 8(a) and 8(b) and pretends to die when the owner pretends to shoot in Figure 8(c). Figures 8(d) to 8(f) show that Cho-Rong-I passes under the owner's legs. As the proposed system obtained the precise position, the augmented dog, Cho-Rong-I, also reacts with the person at the virtual region matching the real coordinates.

6. Conclusion

In this paper, we proposed adaptive sensor fusion methods that compensate for horizontal bias between a laser scanner and video camera for tracking people in real-time system. The usual method using the checkerboard is inconvenient to track people simultaneously because of the manual calibration in a previous research. The proposed system in this paper overcomes the problem using an automatic adaptive sensor fusion method in real-time people tracking. In this application field, the accuracy of compensation is a significant factor for a



(a) (b) (c)

(d) (e) (f)

Figure 8: Demonstration of the virtual pet system.

real-time system. In order to match images between the video camera and laser scanner, we propose the algorithm to merge the laser scanner and video camera simultaneously by capturing the position of the candidate region and the cluster of the laser scan datum. To evaluate the performance of the proposed system, we employed it for tracking two people and then applied it in an augmented reality where one person can interact with a virtual pet. These results show that the proposed system can be successfully employed to obtain the peoples' position by automatic sensor fusion. To enhance the proposed system, variations of the other axis should also be considered. We are currently improving our system to include the uncertainty variations of the sensor position.

Acknowledgments

This work was supported by the Basic Science Research Program through a Grant from the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2012-0005793 and 2011–0013776). This work was also supported by the IT R&D Program of MKE/KEIT "Development of learner-participatory and interactive 3D virtual learning contents technology" (10039165) and the NAP (National Agenda Project) of Korea Research Council of Fundamental Science and Technology.

References

[1] J. Cui, H. Zha, H. Zhao, and R. Shibasaki, "Multi-modal tracking of people using laser scanners and video camera," *Image and Vision Computing*, vol. 26, no. 2, pp. 240–252, 2008.

[2] Y. Bok, Y. Hwang, and I. S. Kweon, "Accurate motion estimation and high-precision 3D reconstruction by sensor fusion," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '07)*, pp. 4721–4726, April 2007.

[3] Q. Zhang and R. Pless, "Extrinsic calibration of a camera and laser range finder (improves camera calibration)," in *Proceedings of the International Conference on Intelligent Robots and Systems (IROS '04)*, pp. 2301–2306, October 2004.

[4] R. Y. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," *IEEE Journal of Robotics and Automation*, vol. 3, no. 4, pp. 323–344, 1987.

[5] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," in *Proceedings of the 7th IEEE International Conference on Computer Vision (ICCV '99)*, pp. 666–673, September 1999.

[6] C. Mei and P. Rives, "Calibration between a central catadioptric camera and a laser range finder for robotic applications," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '06)*, pp. 532–537, May 2006.

[7] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '99)*, pp. 246–252, June 1999.

[8] Y. Tian, A. Senior, and M. Lu, "Robust and efficient foreground analysis in complex surveillance videos," *Machine Vision and Applications*, vol. 23, no. 5, pp. 967–983, 2012.

[9] J. Choi, Y. Cho, K. Cho, S. Bae, and H. Yang, "A view-based multiple objects tracking and human action recognition for interactive virtual environments," *International Journal of Virtual Reality*, vol. 5, no. 3, pp. 1–6, 2006.

[10] P. Ryan, "Count the steps to motivate walking," *Functional U*, vol. 3, no. 3, pp. 12–16, 2005.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 314028, 7 pages
http://dx.doi.org/10.1155/2013/314028

Research Article

The Display System of a Patient's History Using the RFID and Linux

Soo Young Ye[1] and Heung-kuk Jo[2]

[1] *Department of Radiological Science, Catholic University of Pusan, 57 Oryundae-ro, Geumjeong-gu, Busan 609-757, Republic of Korea*
[2] *Division of Computer Information Engineering, Dongseo University, Jurae-dong, Sasang-gu, San 69-1, UIT 106, Busan 617-716, Republic of Korea*

Correspondence should be addressed to Heung-kuk Jo; hkjo@gdsu.dongseo.ac.kr

Received 17 October 2012; Revised 17 January 2013; Accepted 20 January 2013

Academic Editor: Tai-hoon Kim

It is important for doctors to see a patient's medical history when the patient visits the hospital. Currently, doctors research and compare information on their patient's condition through the patient or medical history chart. This method can cause confusion when the patient has several namesakes which can lead to medical errors and malpractice. To prevent such accidents, doctors should have access to the accurate history of their patients through an independent recognition system at any time or any place. Examples of such independent recognition systems are the RFID system, a compacted embedded system, and RFID reader and a TFT-LCD monitor with a unified single system. In this study, to record a patient's medical history, an RFID monitoring system with a carrier frequency of 125 kHz a PXA255 ARM chip, and a 64 Mbyte SDRAM was configured and operated as an embedded Linux system. These two systems were interlocked to read the patient's arm tags. The interlocked system can obtain patients' medical history from a database and display it on a monitor. The advantage of this system is that it does not need a computer and can be implemented with only a hand-held device. Its performance analysis chart shows the result from the experimental device and the results shown with photographs.

1. Introduction

Doctors and nurses need a mobile conversion system in a ubiquitous environment in a hospital to see their patient's personal and medical histories [1–4]. An RFID system is first required, however, for recognition of the patient, as is a loaded embedded Linux system for data monitoring. The RFID tag system is essential for patients who wear a tag, as it recognizes a patient's medical history without error [1–4]. The existing real-time embedded Linux operating system is more advanced than the previous system and supports more features, and through its current source, its development speed is fast increasing [5–9]. Judging from these facts, a mobile ability system is required for the design of a patient monitoring system. Its development cost is also lower. An integrated system for the monitoring of the patient's history must be capable of real-time behavior. Moreover, immediately needed tag information for monitoring requires a fast-booting system, and one of the essential requirements for mobile conversion is a low-power mobile system [10–17]. In this study, to monitor RFID data in real time, the RFID system used a 125 kHz carrier wave with the EM4095 and an embedded Linux operating system with a 400 MHz PXA255 ARM RISC chip, a 512 Kbyte Boot Flash, and a 64 Mbyte SDRAM. Moreover, the system was configured to use a NAND Flash [9]. Given the overall system configuration and operating conditions, CC1020 was embedded in the RFID system using a wireless communication system [18–26]. The extended Linux system boot process and the overall boot loader, kernel, and application processes were operated in an orderly manner. In conclusion, the performance was analyzed and shown with pictures of the experimental systems.

```
si::sysinit:/etc/rc.d/rc.sysinit
...
...
lc:0123456:wait:/etc/rc.d/rc.local
l0:0:wait:/etc/rc.d/rc 0
l1:1:wait:/etc/rc.d/rc 1
l2:2:wait:/etc/rc.d/rc 2
l3:3:wait:/etc/rc.d/rc 3
l4:4:wait:/etc/rc.d/rc 4
l5:5:wait:/etc/rc.d/rc 5
l6:6:wait:/etc/rc.d/rc 6
    cp:0123456:wait:/etc/rc.d/rc.app // added this line
```

Algorithm 1: /etc/inittab files correction.

Table 1: A class of runlevel.

Runlevel 0	Halt (service about system close)
Runlevel 1	Single user mode
Runlevel 2	Multiuser without NFS (multiuser mode without NFS server)
Runlevel 3	Multiuser (multiuser mode)
Runlevel 4	Not used
Runlevel 5	X11 (multiuser mode of login display using X11)
Runlevel 6	Reboot (service about rebooting)

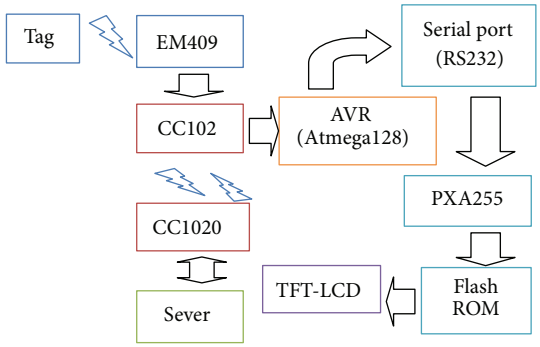


Figure 1: Block diagram of the overall system.

2. System Configuration

2.1. System Concept. In Figure 1, tag information is sent to the 125 kHz RFID reader. The received tag data will be sent to the server using the wireless CC1020 module with a 433 kHz frequency [27–29]. The saved data from the server contains the corresponding patient medical history, and the tag information and corresponding data are retransmitted via a wireless module. After such data are received using RS232 communication, they will be transmitted to the embedded system that is equipped with the PAX255 ARM Chip. Based on the QT written monitoring program, the embedded system outputs a history of the patient's medical history to a TFT-LCD.

Table 2: cp script.

Name	Runlevel	Action	Command
cp	0123456	wait	/etc/rc.d/rc.app

2.2. System Configuration

2.2.1. RFID System. The RFID system used an EM4095 general purpose chip that was produced by the SWATCH group.

It is a CMOS device that uses the 125 kHz frequency band. The tag in the RFID system was EM4100 and will operate only in the Read-Only mode. Using EM4095, the reader system was set up to operate in the Read-Only mode. Figure 2 shows the circuit configuration of the RFID system.

L2 in Figure 2 is the antenna that is attached to the reader. Tag data are entered on the antenna of the EM4095 13-pin output. To input the data to the CC1020, Atmega128 MICOM was used.

2.2.2. Wireless Communication System with CC1020. Figure 3 shows the circuit that used the CC1020 chips. One of the important characteristics is shown in the schematic of the PLL circuit using fixed and variable frequencies by synchronous phase and LC filters.

The frequency of a particular band pass was used. CC1020 was needed to set the behavior of the PSEL, PCLK, PDI, and PDO pins and for a bidirectional AVRData interface. The DIO pins DCLK pin connection was required. An almost perfect CC1020 is shown in Figure 4 with synchronous serial interface settings, using the MCU.

This interface can be divided into two parts. First, the internal register has to install a CC1020 chip that supports the implementation of a communication system. After the completion of the second set, the actual data are used for transmission and receipt. The interface circuit in Figure 4 explains this process.

2.2.3. Embedded Linux System. EZBOOT was used for the embedded Linux system that was manufactured by Falinux. Its main feature is its ability to memorize, read, and write.

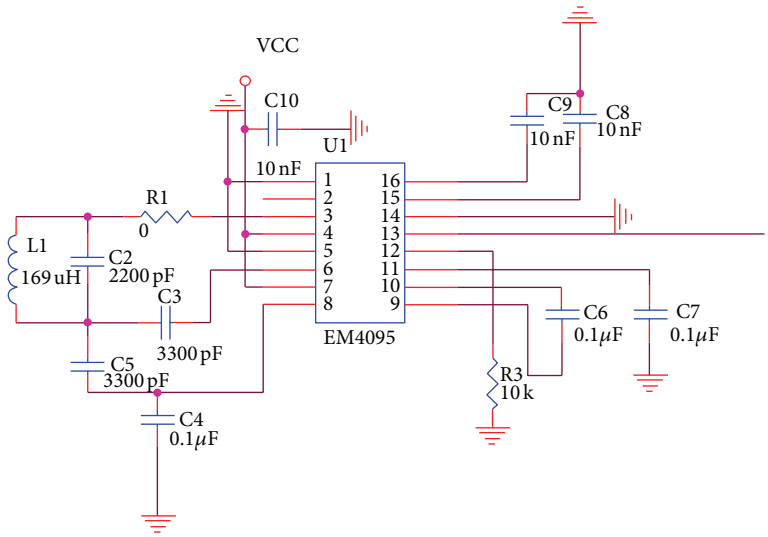


Figure 2: Circuit of RFID system.

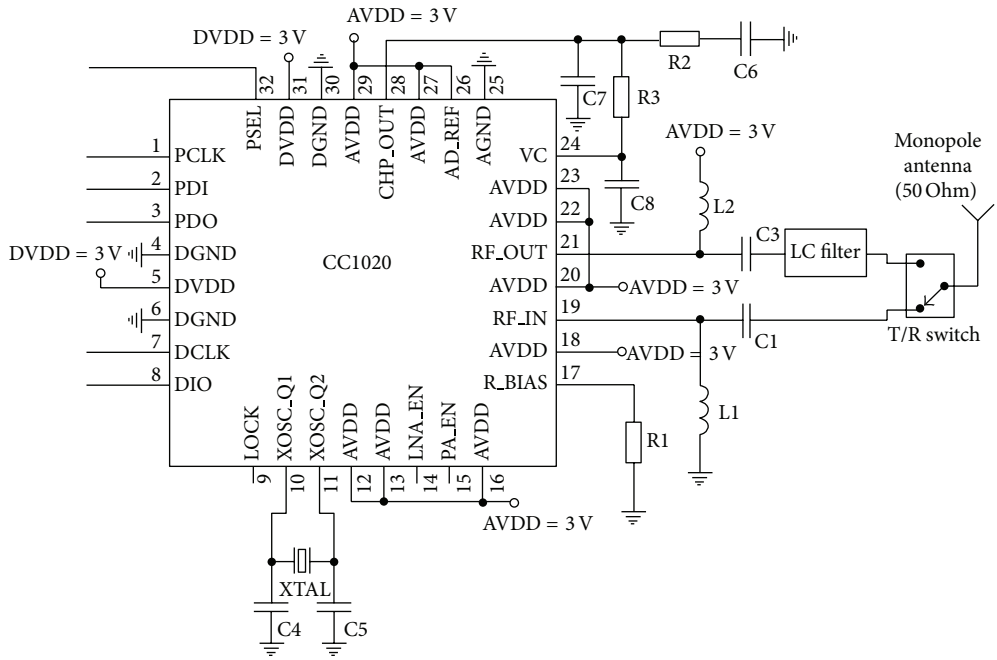


Figure 3: Circuit of transmitter and receiver part of CC1020.

It uses ZMODEM to download images and flash memory to read and write. To display the history of the patient information system on the LCD, the embedded Linux system must be initialized. The initialization process should be performed as shown in Figure 5.

At the moment of the initialization of the embedded Linux system, the Boot loader must be executed in 512 Kbytes of the BIOS Boot Flash within sector 0, cylinder 0. The Boot loader will first initialize the hardware for the system to operate, and the CPU will make the access possible to the SDRAM. As shown in Figure 6, the kernel and RAMDISK images are stored in the 64 Mbyte NAND Flash and will be saved in the 64 Mbyte SDRAM.

Before the kernel is performed for the i386 Linux-2.4.19 to make it apt for ARM usage, ARM patching must be performed. Due to the use of a CPU as a PXA255 chipset, the kernels with the performed XScale must be loaded to the 64 M NAND Flash in the EZBOOT setup mode. The Linux kernel and arch/*/kernel/head.S, which enable the C code to perform the operation, were programmed at the entry point. Moreover, the init/main.c has a start kernel ( ) function, and such function of arch_init, trap_init, is

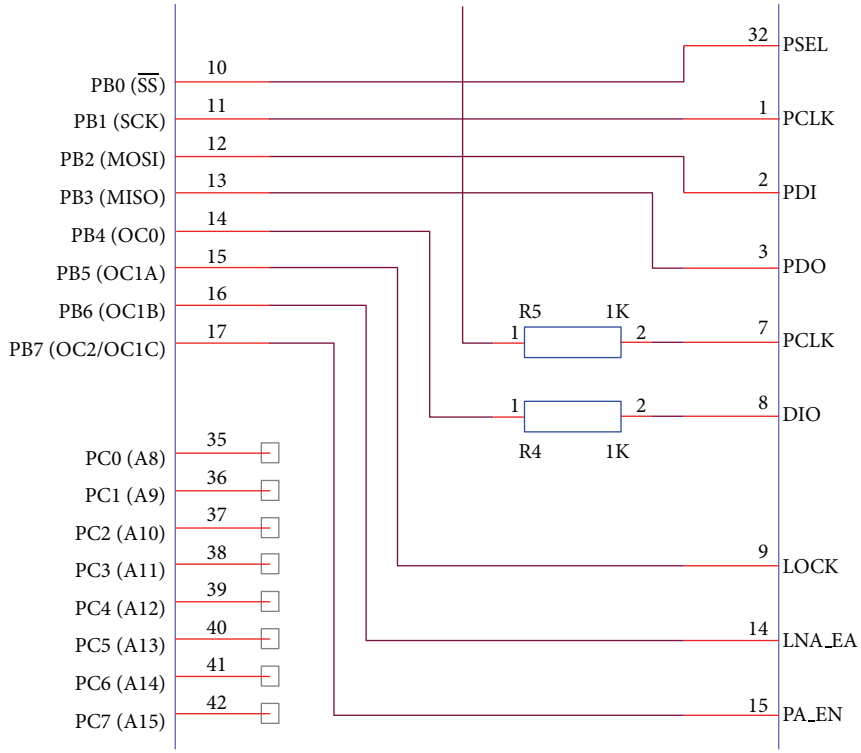


Figure 4: Interface between Atmega128 and CC1020.

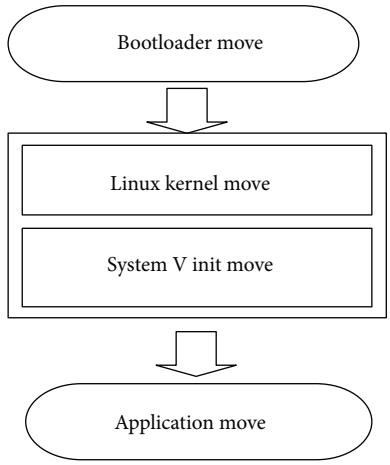


Figure 5: Flowchart for initial process of the embedded Linux system.

an architecture-dependent initialization function. Each part of the Linux kernel has an initialization function. The start kernel ( ) function at the end of the kernel thread (init,. . .), as a function of the PID, will create a single kernel thread, and the start kernel ( ) function itself is an idle task. Moreover, the init kernel thread will mount the root file system in the Read-Only mode. At this point, because the embedded Linux system does not have a hard disk, it must load the root file system to the Ram disk. When the Ram disk that serves as the root file system is loaded, the System V init is finally performed. The last kernel performance process in the exec/sbin/init is called the System V init program. It has the same functions and performance as the DOS program of the old, which had autoexec.bat (the Batch file that enabled automatic operation). The System V init and /etc./inittab file perform according to the contents of the inittab file, and the information contents vary depending on such performance. Table 1 shows the Runlevel per performance type, which groups the services into different categories.

In the embedded system, a specific user application, the booting must be carried out simultaneously. Therefore, the RFID tag data must be sent from the system and outputted to the TFT-LCD window immediately after the board booting is completed. Thus, the /etc./inittab file shown in Algorithm 1 was revised.

The Runlevel of the embedded Linux systems is basically 3, l3:3:wait:/etc/rc.d/rc 3. The execution sequence of inittab is rc.sysinit>>rc.local>>rc3.d. As shown in Table 2, an executable script was added, and in the directory of etc./rc.d that was executed with rc.local, the rc.app application script was made. A brief description of the script follows.

3. Experiment and Performance Analysis

Figure 7 shows the future test of RFID systems and the embedded Linux systems in this project. Figure 7(a) shows a reading from the RFID tag system, which indicates that the TFT-LCD can display related data. The PXA255 chip, 64 Mbyte SDRAM, and NAND Flash are configured to be compatible with the embedded Linux system. Moreover,

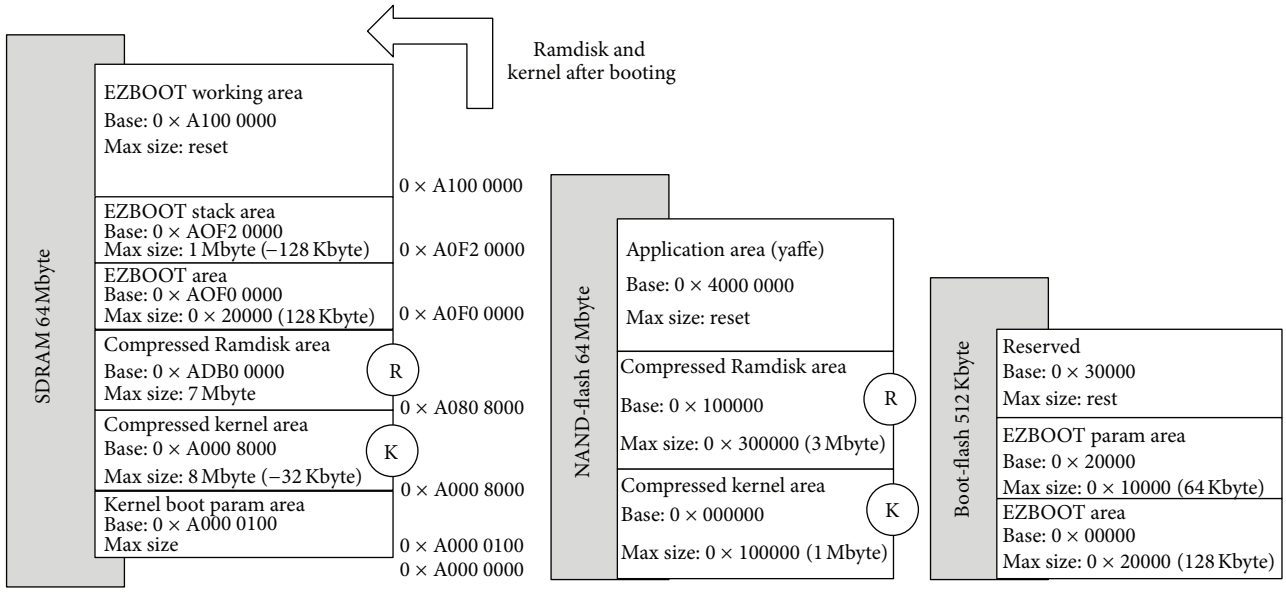
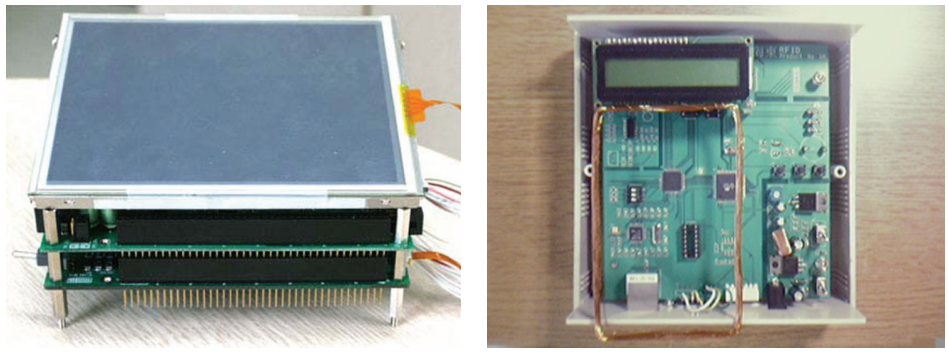


Figure 6: Boot loader memory map.



(a) Embedded LINUX system (b) RFID system

Figure 7: Experiment system.

Figure 7(b) shows the capability of the 125 kHz RFID system and server, which are configured to communicate with a wireless communication module, the CC1020 system, to read tag information.

Figure 8 tag shows that the RFID system has read the data. The output of the TFT-LCD window shows the experiment results of the embedded Linux system.

As shown in the picture, if the tag has not made contact with the reader, a basic screen output will be shown like the picture shown in Figure 8(a).

If contact has been made between the reader and the tag, the data are immediately recognized and completed. These data will complete the communication through the server, and the output image will be displayed by the embedded systems of the TFT-LCD screen, as shown in Figure 8(b).

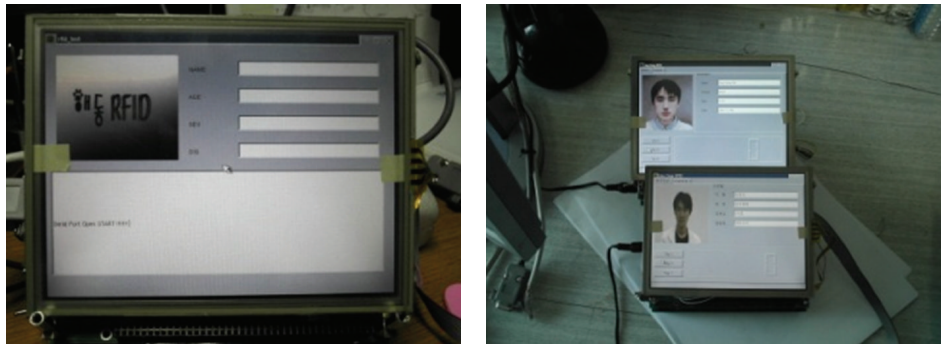
Table 3 shows the performance of the experiment device. The CC1020 transmission of the transferred data among the tag IDs involved removal of the starting byte and the stop byte of the tag ID. After only 8 bytes were removed, the information was transmitted and received. The experiment was repeated 100 times using the system, and no error occurred during the transmission.

4. Conclusion

In this study, through a digital wireless communication device, tag information data were transmitted between the client server and the server with an RFID system. Concerning the verified data, a system that could monitor information in real time was developed: the embedded Linux system with a PXA255 ARM chip. The RFID system developed a multichannel, unidirectional, and bidirectional transmission/receiving system using the ISM band. Its distinctive characteristics are low-power consumption, high sensitivity, a small chip, and

Table 3: System performance.

	Number of data byte	Baud rate	Method of communication	
RFID	10 Byte	9600 bps	Wireless	
Embedded Linux	Depend on DB	9600 bps	Wired	
CC1020 transmitter	Depend on DB	57600 bps	Wireless	Transmit range: 50 m
CC1020 receiver	Depend on DB	57600 bps	Wireless	



(a) Initial sense (b) Sense of tag data on monitoring system

Figure 8: Display on the TFT-LCD monitoring.

fewer external components. Therefore, the system showed the possibility of low-power usage and reduced production costs and sizes. Embedded Linux systems are integrated with the RFID tag System, and information must be outputted on a TFT-LCD window and must internally use EZBOOT with 2.4 kernels. This system was developed to simultaneously self-boot and run the QT application, so that it can monitor RFID data. The experiment was performed inside a walled building; and within 50 m, tag data were transmitted 100 times. There were no errors during the tag data monitoring. With this result, this system showed potential application in hospitals.

Acknowledgment

This research was supported by a research program of Dongseo University's Ubiquitous Appliance Regional Innovation Center supported by the Grants from the Ministry of Knowledge Economy of the Korean government and Busan Metropolitan City (no. B0008352).

References

[1] T. P. Quoc, M. C. Kim, H. K. Lee, and K. H. Eom, "Wireless sensor network apply for the blind U-bus system," *International Journal of u- and e- Service, Science and Technology*, vol. 3, no. 3, 2010.

[2] Y. S. Kim, "A network-based telemedicine service," *International Journal of Energy, Information and Communications*, vol. 2, no. 1, 2011.

[3] P. T. B. Hue, S. Wohlgemuth, and I. Echizen, "An experimental evaluation for a new column-level access control mechanism health record systems," *International Journal of u- and e-Service, Science and Technology*, vol. 4, no. 3, 2011.

[4] H. K. Jo, *Production Manual of RFID System Using U2270B*, Hong-Neng Publish, 2008.

[5] H. J. Lee and H. K. Jo, *Security and System of RFID*, Com-One Media, 2009.

[6] H. K. Jo and H. J. Lee, "A relay-transmission of the RFID tag ID over the wireless and TCP/IP with a security agent," in *Agent and Multi-Agent Systems: Technologies and Applications*, vol. 4496 of *Lecture Notes in Computer Science*, pp. 918–927, Springer, New York, NY, USA, 2007.

[7] Y. H. Kim and H. K. Jo, "Patient information display system in hospital using RFID," in *Proceedings of the International Conference on Hybrid Information Technology (ICHIT '09)*, pp. 397–400, 2009.

[8] W. Stalling, *Data and Computer Communications*, Prentice Hall, New York, NY, USA, 8th edition, 2007.

[9] L. Garcia and I. Widjaja, *Communication Network: Fundamental Concepts and Architectures*, McGraw-Hill, New York, NY, USA, 2001.

[10] M. J. Donahoo and K. L. Calvert, "TCP/IP Socket Programming version C," 2001.

[11] S. Yun, "TCP/IP Socket Programming," 2003.

[12] D.-Y. Yun, "AVR ATmega128 Master," 2004.

[13] Technical Datasheet V1.32, http://www.wiznet.co.kr/.

[14] K. Finkenzeller, *RFID Handbook*, John Wiley & Sons, New York, NY, USA, 1999.

[15] Jonathan Collins, "FTC asks RFID users to self-regulate," *RFID Journal*, 2005.

[16] E. Kaasinen, *User acceptance of mobile services-value. Ease of use, trust and ease of adoption [Doctoral dissertation]*, VTT Publication, 2005, VTT 566.

[17] V. Deolalikar, M. Mesarina, J. Recker, and S. Pradon, "Perturbative time and frequency allocations for RFID reader networks," in *Proceedings of the international conference on Emerging*

Directions in Embedded and Ubiquitous Computing (EUC '06), 2006.

[18] K.-H. Lee, Y.-Y. An, H.-D. Park, and Y.-Z. Cho, "A data-centric self-organization schema for energy-efficient wireless sensor networks," in *Proceedings of the International Conference on Emerging Directions in Embedded and Ubiquitous Computing (EUC'06)*, Seoul, South Korea, 2006.

[19] S. C. Kim, S. S. Yeo, and S. K. Kim, "MARP: mobile agent for RFID privacy protection," in *Smart Card Research and Advanced Applications*, vol. 3928 of *Lecture Notes in Computer Science*, pp. 300–312, 2006.

[20] J. G. Choi, *Visual C#. NET, 2005 2nd Edition Project*, Young-Jin Publish, 2007.

[21] I. G. Kim, G. S. Kim, S. J. Ha, and Y. J. Kim, "Programming of C#. NET," Inter-vision, 2007.

[22] Institute Technology Micom World, *Theory and Experiment Using AVR, (ATmega128)*, Micom-World, 2004.

[23] http://www.atmel.com/avr/.

[24] http://www.emmicroelectronic.com/.

[25] http://www.Firmtech.co.kr/.

[26] http://www.realtek.com.tw/.

[27] Chipcon Tech. Support, "User Manual Rev. 2.0," Chip AS, 2004, http://www.chipcon.com/.

[28] Chipcon Tech. Support, "User Manual Rev. 2.3," Chip AS, 2003.

[29] Chipcon Tech. Support, "CC1020 Datasheet (rev. 1.5)," Chip AS.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 460641, 15 pages
http://dx.doi.org/10.1155/2013/460641

Research Article

Optimizing Classification Decision Trees by Using Weighted Naïve Bayes Predictors to Reduce the Imbalanced Class Problem in Wireless Sensor Network

Hang Yang,[1] Simon Fong,[1] Raymond Wong,[2] and Guangmin Sun[3]

[1] *Department of Computer and Information Science, University of Macau, Taipa, Macau*
[2] *School of Computer Science and Engineering, University of New South Wales, Sydney, NSW 2052, Australia*
[3] *Department of Electronic Engineering, Beijing University of Technology, Beijing 100022, China*

Correspondence should be addressed to Simon Fong; ccfong@umac.mo

Received 6 October 2012; Accepted 18 October 2012

Academic Editor: Sabah Mohammed

Standard classification algorithms are often inaccurate when used in a wireless sensor network (WSN), where the observed data occur in imbalanced classes. The imbalanced data classification problem occurs when the number of samples in one class, usually the class of interest, is much lower than the number in the other classes. Many classification models have been studied in the data-mining research community. However, they all assume that the input data are stationary and bounded in size, so that resampling techniques and postadjustment by measuring the classification cost can be applied. In this paper, we devise a new scheme that extends a popular stream classification algorithm to the analysis of WSNs for reducing the adverse effects of the imbalanced class in the data. This new scheme is resource light at the algorithm level and does not require any data preprocessing. It uses weighted naïve Bayes predictors at the decision tree leaves to effectively reduce the impact of imbalanced classes. Experiments show that our modified algorithm outperforms the original stream classification algorithm.

1. Introduction

A wireless sensor network (WSN) is a distributed platform that collects data over a broad area. It has a wide variety of practical military, medical, and industrial applications [1]. The brain of a WSN is usually a decision-making algorithm that is capable of correctly mapping a set of newly collected observations from the sensors to one or more predefined categories. It uses a machine-learning algorithm to recall the classification of old data and classify the new data accordingly. There is no shortage of machine-learning algorithms available for decision making in WSNs [2, 3]. However, the imbalanced classification is a common problem. This problem occurs when the classifier algorithm is trained with a dataset in which one class has only a few samples and there are a disproportionally large number of samples in the other classes. This kind of imbalanced data causes classifiers to be overfitted (i.e., produce redundant rules that describe duplicate or meaningless concepts) and as a result perform poorly, particularly in the identification of the minority class. In WSN applications, these rare minority classes are often critical. Some WSN examples include, but are not limited to, transaction fraud detection, machine fault monitoring, environmental anomalies, atypical medical conditions, and abnormal habitual behaviors—situations where the class of interest is a small sample of unusual readings. Studies [4] have shown that using standard classification algorithms to analyze these imbalanced class distributions leads to poor performance. An imbalanced class problem may have another implication in WSN where it could be a symptom of producing traffic "hot-spot" in WSN. The energy consumption in the sensors may become imbalanced too, which leads to premature drainout for some local nodes. Some solution [5] has been proposed to better cluster the nodes and traffics although it is aimed at the energy level.

Most of the standard classification algorithms assume that training examples are evenly distributed among different classes. In practical applications where this was known to be

untrue, researchers addressed the problem by either manipulating the training data or adjusting the misclassification costs. Resizing the training datasets is a common strategy that attempts to downsize the majority class and oversamples the minority class. Many variants of this strategy have been proposed [6–8]. A second strategy is to adjust the costs of misclassification errors to be biased against or in favor of the majority and minority classes, respectively. Using the feedback from the altered error information, researchers then [9, 10] fine-tune their cost-sensitive classifiers and postprune the decision trees in the hope of establishing a balanced treatment of each class in the new imbalanced data collected by the network.

The authors of this paper argue that replacing the traditional classifier with an optimized stream classifier is another effective solution. As mentioned above, the current techniques for dealing with imbalanced data require additional data preprocessing or feedback learning and pruning of a trained decision tree. Though they may be useful in minimizing the impact of imbalanced data, these pre- and postprocessing mechanisms require working through a whole database and their operations incur certain overheads in the data-mining environment, which may not be favorable in a WSN. In a wireless sensor network, data mining is done in real time with a compact device with limited memory and processing power called a sink, and most importantly the incoming data for classification training and testing are streaming in nature. These data streams are nonstationary data that may only be read one time at the intermediate nodes of a sensor network and are then forgotten. Furthermore, these nodes may be required to perform real-time classification as the data flows along the WSN. Fast prediction results with satisfactory accuracy must be propagated from node to node. In this dynamic environment, techniques based on data storage and feedback style after learning cannot be used to correct imbalanced data. On the other hand, it has been demonstrated that a stream classifier is a good candidate for WSN applications [11].

The contribution of this paper is a set of simple modifications that optimize an existing stream classification algorithm called Very Fast Decision Tree to handle the imbalanced class problem at the algorithmic level. One important extension is the use of weighted naïve Bayes predictors installed at the decision tree leaves. The assigned weights have the effect of countering the "biases" that are introduced by the problems of imbalanced class found in imperfect WSN data. The paper is organized as follows. Section 2 describes in detail the modifications that tackle the imbalanced class problem. In Section 3 a range of experiments is described, the "biased" datasets with imbalanced classes are introduced and the experimental results are discussed. Section 4 concludes the paper.

2. Optimizing the Very Fast Decision Tree

2.1. Motivation and Overview. Three special modifications are proposed to enhance the Very Fast Decision Tree (VFDT) algorithm. These modifications are embedded in line with the codes that implement the classification logic of the stream classifier. The modifications to reduce the imbalanced class problem are made in four phases: the training phase, where new nodes are created if the statistical criteria established in the learning from the labeled samples phase are met; the prepruning phase, in which the qualified nodes and branches are tested to see whether they can indeed improve the prediction accuracy (before they are added to the decision tree); the prediction phase where unseen samples are being categorized to predefined classes; and the pruning phase which uses the functional tree leaf [12]. The modifications are in the forms of simple computation and conditional checks that do not incur heavy resource consumption at the sensor nodes.

The improved version of VFDT is generally called the Optimized Very Fast Decision Tree with Functional Leaves (OVFDT-FL). Our previous work [13] shows that the OVFDT-FL prototype can classify data streams with the maximum possible accuracy with the minimum tree size. In this paper, OVFDT-FL is tested with imbalanced class data. The design of OVFDT-FL is given as follows.

OVFDT, which is based on the original VFDT design, is implemented using a test-then-train approach for classifying a continuously arriving data stream, even $N \to \infty$ where N is the total number of training instances as shown in Figure 1. The whole test-then-train process is synchronized so that when the data stream arrives one segment at a time, the decision tree is tested first for prediction output and training (which is also known as updating) of the decision tree model then occurs incrementally. Suppose that X is a vector of d attributes and k is the number of classes included in the data streams. When a new data sample (X, y_k) arrives, it travels from the root of the decision tree to an existing leaf via the current decision tree structure, provided that the root existed initially. Otherwise, a heuristic function is used to construct a tree model with a single root node using the procedure shown in Pseudocode 1. Suppose that a decision tree model HT can give a prediction to a class $y_k^{'}$ according to the functional tree leaf $\mathcal{F}$, where $\mathrm{HT}(X) \to y_k^{'}$. Comparing the predicted class $y_k^{'}$ to the actual class y_k, the statistics of the true, C_T, and false, C_F, predictions are updated immediately. Meanwhile, the sufficient statistics n_{ijk}, which are a count of the attribute x_i with value j that belong to class y_k, are updated for each node. This series of actions is called the testing phase.

The training phase immediately follows the testing phase. Node-splitting estimation is used to initially decide if HT should be updated or not, depending on the number of samples received that can potentially be represented by additional underlying rules in the decision tree. In principle, the node-splitting estimation should apply to every single new sample that arrives. However, this would be too resource expensive and would slow down the tree building process. Instead, VFDT proposes a parameter $n_{\min}$ that only carries out the node-splitting estimation when $n_{\min}$ examples have been observed on a leaf. In the node-splitting estimation, the tree model should be updated when a heuristic function $G(\cdot)$ chooses the most appropriate attribute, with the highest heuristic function value

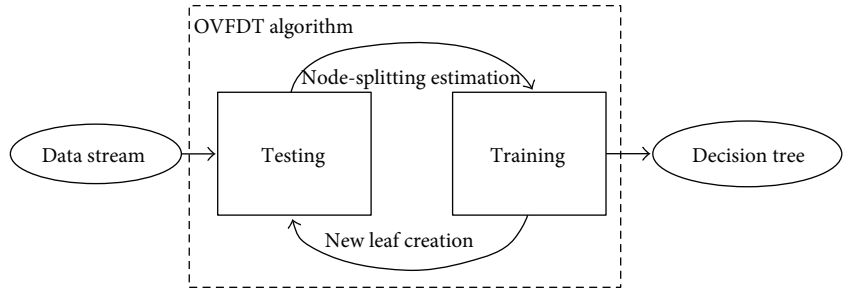


Figure 1: A test-then-train OVFDT workflow.

INPUT:
S: A stream of sample
X: A set of symbolic attributes
$G(\cdot)$: Heuristic function using for node-splitting estimation
δ: One minus the desired probability of choosing a correct attribute at any given node
$n_{\min}$: The minimum number of samples between check node-splitting estimation
$\mathcal{F}$: A functional tree leaf strategy
OUTPUT:
HT: A decision tree

PROCEDURE: initializeHT$(S, X, G(\cdot), \delta, n_{\min})$
(1) Let HT be a tree with a single leaf l (the root). Let $X_l = X \cup \{X_\varnothing\}$
(2) Let G_l $(X_\varnothing)$ be the $G(\cdot)$ obtained by predicting the class in S, according to $\mathcal{F}$.
(3) FOR each class y_k
(4) //y_k *is the class lable y with the kth label*
(5) FOR each value X_{ij}, of attribute $X_i \in X$
(6) Reset OCD: $n_{ijk}(l) = 0$
(7) //$n_{ijk}(l)$ *is the count of attribute with* X_{ij} *and* y_k *at leaf l*
(8) END-FOR
(9) END-FOR
(10) Return HT with a single root

Pseudocode 1: The pseudocode of initializing an OVFDT model.

$G(x_a)$, as a splitting node, according to Hoeffding's bound and the tie-breaking threshold. The heuristic function is implemented as an information gain here. This in situ system of node-splitting estimation constitutes our training phase.

Two modifications are proposed for the training phase of OVFDT to manage imbalanced data classes. The first is to dynamically adjust the tie-breaking function in the splitting-node determination using the mean value of Hoeffding's bound. The growth of the tree is influenced by the mean value of the traffic fluctuation (which was found to correlate with Hoeffding's bound in our previous work) rather than the imbalanced data class. The second modification is to use prepruning to test if the leaf chosen to be split, and therefore increase tree growth, is indeed a valid choice given the imbalanced data class. In this way, we can assume that the expansion of the tree is a result of genuinely accurate predictions. Thus, postpruning on the decision tree is not necessary. Section 2.2 presents the details of the functional leaf strategy for handling the imbalanced data class, and the details of the modifications to the training phase are given in Section 2.3.

2.2. Functional Tree Leaf Prediction in Testing Phase. The sufficient statistics n_{ijk} is an incremental count number stored in each node in the OVFDT. Suppose that a node Node_{ij} in HT is an internal node labeled with attribute x_{ij}. Suppose that k is the number of classes distributed in the training data, where $k \geq 2$. A vector V_{ij} is constructed from the sufficient statistics n_{ijk} in Node_{ij} such that $V_{ij} = \{n_{ij1}, n_{ij2}, \ldots, n_{ijk}\}$. V_{ij} is the observed class distribution (OCD) vector of Node_{ij}. OCD stores the count of the distributed class at each tree node in OVFDT. It helps to keep track of the occurrences of the instances of each attribute.

For the actual classification, OVFDT uses $\text{HT}(X) \rightarrow y_k^{'}$ to predict the class label when a new sample (X, y) arrives. The predictions are made according to the OCD in the leaves, which is called the functional tree leaf $\mathcal{F}$. Originally, in VFDT the prediction used only the majority class functional tree leaf

PROCEDURE: traverseHT(S, HT, $\mathcal{F}$)
(1) Sort S from the root to a leaf by HT. Update OCD in each node: $n_{ijk}(l)$ ++
(2) Switch ($\mathcal{F}$)
(3) Case $\mathcal{F}^{\mathrm{MC}}$: predict the class $y_k^{'}$ with max $n_{ijk}(l)$
(4) Case $\mathcal{F}^{\mathrm{NB}}$: predict the class $y_k^{'}$ with max NB prob.
(5) Case $\mathcal{F}^{\mathrm{WNB}}$: predict the class $y_k^{'}$ with max WNB prob.
(6) Case $\mathcal{F}^{\mathrm{Adaptive}}$: predict the class $y_k^{'}$ using $\mathcal{F}$ with $\mathrm{Error}_{\min}$
(7) IF $y_k^{'}$ equals to the actual class label in S, THEN C_T++
(8) ELSE C_F++
(9) $\Delta C = C_T - C_F$
(10) Return ΔC

Pseudocode 2: The pseudocode of OVFDT testing phase.

$\mathcal{F}^{\mathrm{MC}}$. The majority class only considers the counts of the class distribution, but not the decisions based on combinations of attributes. The naïve Bayes functional tree leaf $\mathcal{F}^{\mathrm{NB}}$ was proposed to compute the conditional probabilities of the attribute values given a class at the tree leaves by naïve Bayes. As a result, the prediction at the leaf is refined by the consideration of the probabilities of each attribute. To handle imbalanced class distribution in a stream, a weighted naïve Bayes functional tree leaf $\mathcal{F}^{\mathrm{WNB}}$ and an adaptive functional tree leaf $\mathcal{F}^{\mathrm{Adaptive}}$ are proposed in the paper.

2.2.1. Majority Class Functional Tree Leaf. In the ODC vector, the majority class functional tree Leaf $\mathcal{F}^{\mathrm{MC}}$ chooses the class with the maximum distribution as the predictive class in a leaf, where $\mathcal{F}^{\mathrm{MC}}$: $\arg\max f = \{n_{i,j,1}, n_{i,j,2}, \ldots, n_{i,j,r}, \ldots, n_{i,j,k}\}$, and where $0 < r < k$.

2.2.2. Naïve Bayes Functional Tree Leaf. In the OCD vector $V_{i,j} = \{n_{i,j,1}, n_{i,j,2}, \ldots, n_{i,j,r}, \ldots, n_{i,j,k}\}$, where r is the number of observed classes and $0 < r < k$, the naïve Bayes functional tree leaf $\mathcal{F}^{\mathrm{NB}}$ chooses the class with the maximum possibility, as computed by the naïve Bayes, as the predictive class in a leaf. $n_{i,j,r}$ is updated to $n_{i,j,r}^{'}$ by the naïve Bayes function such that $n_{i,j,r}^{'} = \mathrm{P}(X \mid C_f) \cdot \mathrm{P}(C_f)/\mathrm{P}(X)$, where X is the new arrival instance. Hence, the prediction class is $\mathcal{F}^{\mathrm{NB}}$: $\arg\max i = \{n_{i,j,1}^{'}, n_{i,j,2}^{'}, \ldots, n_{i,j,r}^{'}, \ldots, n_{i,j,k}^{'}\}$.

2.2.3. Weighted Naïve Bayes Functional Tree Leaf. In the OCD vector $V_{i,j} = \{n_{i,j,1}, n_{i,j,2}, \ldots, n_{i,j,r}, \ldots, n_{i,j,k}\}$, where k is the number of observed classes and $0 < r < k$, the weighted naïve Bayes functional tree leaf $\mathcal{F}^{\mathrm{WNB}}$ chooses the class with the maximum possibility, as computed by the weighted naïve Bayes, as the predictive class in a leaf. $n_{i,j,r}$ is updated to $n_{i,j,r}^{'}$ by the weighted naïve Bayes function such that $n_{i,j,r}^{'} = \omega_r \cdot \mathrm{P}(X \mid C_f) \cdot \mathrm{P}(C_f)/\mathrm{P}(X)$, where X is the new arrival instance, and the weight is the probability of class i distribution amongst all the observed samples such that $\omega_r = \prod_{r=1}^{k}(v_r/\sum_{r=1}^{k} v_r)$, where $n_{i,j,r}$ is the count of class r. Hence, the prediction class is $\mathcal{F}^{\mathrm{WNB}}$: $\arg\max f = \{n_{i,j,1}^{'}, n_{i,j,2}^{'}, \ldots, n_{i,j,r}^{'}, \ldots, n_{i,j,k}^{'}\}$.

2.2.4. Adaptive Functional Tree Leaf. In a leaf, suppose that $V_{\mathcal{F}^{\mathrm{MC}}}$ is the observed class distribution vector with the majority class functional tree leaf $\mathcal{F}^{\mathrm{MC}}$, suppose that $V_{\mathcal{F}^{\mathrm{NB}}}$ is the observed class distribution vector with the naïve Bayes functional tree leaf $\mathcal{F}^{\mathrm{NB}}$, and suppose that $V_{\mathcal{F}^{\mathrm{WVB}}}$ is the observed class distribution vector with the weighted naïve Bayes functional tree leaf $\mathcal{F}^{\mathrm{WNB}}$. Suppose that y is the true class of a new instance X. Suppose that $E_{\mathcal{F}}$ is the prediction error rate using a functional tree leaf $\mathcal{F}$. $E_{\mathcal{F}}$ is calculated by the average $E = \mathrm{error}_i/n$, where n is the number of examples and error_i is the number of examples mispredicted using $\mathcal{F}$. The adaptive functional tree leaf chooses the class with the minimum error rate predicted by the other three strategies, where $\mathcal{F}^{\mathrm{Adaptive}}$: $\arg\min \mathcal{F} = \{E_{\mathcal{F}^{\mathrm{MC}}}, E_{\mathcal{F}^{\mathrm{NB}}}, E_{\mathcal{F}^{\mathrm{WVB}}}\}$.

According to the functional tree leaf strategy, the current HT sorts a newly arrived sample (X, y_k) from the root to a predicted leaf $y_k^{'}$. Comparing the predicted class $y_k^{'}$ to the actual class y_k, the statistics of truly C_T and falsely C_F prediction are updated immediately. C_T and C_F are used in the model-training phase. Pseudocode 2 is a flowchart of the modified testing phase.

2.3. Dynamic Splitting Test and Prepruning in the Training Phase. The node-splitting control is modified to use a dynamic tie-breaking threshold τ, which restricts the attribute splitting at a decision node. The τ parameter traditionally is preconfigured with a default value defined by the user. The optimal value is usually not known until all of the possibilities in an experiment have been tried. Longitudinal testing of different values in advance is certainly not favorable in real-time applications. Instead, we assign a dynamic tie threshold, equal to the dynamic mean of the HB value at each pass of stream data, as the splitting threshold, which controls the node splitting during the tree-building process. Tie breaking that occurs close to the HB mean can effectively narrow the variance distribution. The HB mean is calculated dynamically whenever new data arrives and the HB value is updated.

The estimation of splits and ties is only executed once for every $n_{\min}$ (a user-supplied value) sample that arrives at a leaf. Instead of a pre-configured tie, OVFDT uses an adaptive tie that is calculated by incremental computing. At the ith node-splitting estimation, Hoeffding's bound ε estimates whether there are sufficient statistics from a large enough sample size to split a new node, which corresponds to the leaf l. Let T_l be an adaptive tie corresponding to leaf l, within k estimations seen so far. Suppose that μ_l is a binary variable that takes the value of 1 if ε relates to leaf l and 0 otherwise. T_l is computed by (1). To constrain HB fluctuation, an upper bound T_l^{UPPER} and a lower bound T_l^{LOWER} are proposed in the adaptive tie mechanism. The formulas are presented in (2) and (3) as follows:

$$\mathrm{T}_l = \frac{1}{k}\sum_{i=1}^{k}\mu_l \times \varepsilon_i, \tag{1}$$

$$T_l^{\mathrm{UPPER}} = \arg\max \mathrm{T}_l, \tag{2}$$

$$T_l^{\mathrm{LOWER}} = \arg\min \mathrm{T}_l. \tag{3}$$

For resource-light operations, we propose an error-based prepruning mechanism for the OVFDT, which stops noninformative node splitting before it splits into a new node. The prepruning takes into account both global and local node-splitting errors.

Lemma 1 (Monitoring Global Accuracy). *The model's accuracy varies whenever a node splits and the tree structure is updated. Overall accuracy of a current tree model is monitored during node splitting by comparing the number of correctly and incorrectly predicted samples. The numbers of correctly predicted instances and otherwise are recorded as current global performance indicators. This monitoring allows the determination of global accuracy.*

When a new instance arrives, it will be sorted to a leaf by the current HT structure before the node-splitting estimation. This is the "testing" phase in OVFDT. Suppose that C_T is the number of correctly predicted instances in the current HT and C_F is the number of incorrectly predicted instances. After the ith node-splitting estimation, let ΔC_i be the difference between C_T and C_F, then ΔC_i is computed by (4), which reflects the global accuracy of the current HT prediction on the newly arrived data streams. If $\Delta C_i \geq 0$, the number of correct predictions is no less than the number of incorrect predictions in the current tree structure; otherwise, the current tree graph needs to be updated by node splitting.

Lemma 2 (Monitor Local Accuracy). *The global accuracy can be tracked by comparing the number of correctly predicted samples with the number of incorrectly predicted samples. Likewise, comparing the global accuracy as measured at the current node-splitting estimation with the global accuracy measured at the previous splitting, means that the variation in accuracy is being tracked dynamically. This monitoring allows us to check whether the current node splitting is advantageous at each step by comparing it with the previous step.*

Suppose that $\mathrm{Gain}_{\mathrm{Accu}}$ is the gain in accuracy of the ith and the $(i-1)$th estimations, as calculated in (5), which reflects a local accuracy of changes. If $\mathrm{Gain}_{\mathrm{Accu}}(\mathrm{HT}_i) \geq 0$, the measurement of accuracy at the ith splitting HT structure is no worse than the accuracy at the $(i-1)$th splitting; otherwise, the old tree structure needs to be updated. The splitting estimation is implemented once for every $n_{\min}$ sample that arrives at a leaf. The tree size increases by l when a new node splits. The number of samples that meets the first pruning condition is $(n_{\min} \cdot p)$, where p is the probability of the optimal node splitting calculated in (8). Only one value of p can be chosen at one splitting estimation. The calculation of tree size at estimation i is given in (6). C_T and C_F in the ith splitting estimation give feedback on the tree's current classifying accuracy. By continually comparing this with $(i-1)$th, the pruning maintains the accuracy sequentially. In other words, the optimum result is obtained by comparing the current tree status to its previous status as follws:

$$\Delta C_i = C_T - C_F, \tag{4}$$

$$\mathrm{Gain}_{\mathrm{Accu}}\left(\mathrm{HT}_i\right) = \Delta C_i - \Delta C_{i-1}, \tag{5}$$

$$L_i = L_{i-1} + n_{\min} \cdot p, \tag{6}$$

$$\mathrm{Gain}_{\mathrm{Tree\ Size}}\left(\mathrm{HT}_i\right) = L_i - L_{i-1}, \quad \left(L_0 = 1\right), \tag{7}$$

$$p = \begin{cases} \mathrm{Prob}\left[\Delta G \leq T_l^{\mathrm{LOWER}}\right] \cdot \mathrm{Prob}\left[\Delta C_i < \Delta C_{i-1}\right] & \text{or} \\ \mathrm{Prob}\left[\Delta G \leq T_l^{\mathrm{LOWER}}\right] \cdot \mathrm{Prob}\left[\Delta C_i < 0\right] & \text{or} \\ \mathrm{Prob}\left[T_l^{\mathrm{LOWER}} < \Delta G < T_l^{\mathrm{UPPER}}\right] \cdot \mathrm{Prob}\left[\Delta C_i < \Delta C_{i-1}\right]. \end{cases} \tag{8}$$

Figure 2 shows why our proposed prepruning takes into account both the local and the global accuracy in the incremental pruning. At the ith node-splitting estimation, the difference between correctly and incorrectly predicted classes was ΔC_i, and ΔC_{i+1} at the $i+1$th estimation. $\mathrm{Gain}_{\mathrm{Accu}}(\mathrm{HT}_{i+1})$ was negative, indicating that the local accuracy of $i+1$th estimation was worse than that at the previous node-splitting, although both were on a globally increasing trend. Thus, if accuracy is declining locally, it is necessary to update the HT structure even if accuracy is increasing globally.

The optimal node splitting control consists of a dynamic tie for node splitting and a prepruning mechanism that tries to hold the tree growth in neutral with respect to the imbalanced class distribution. In each node-splitting estimation process, the Hoeffding bound (HB) value that relates to leaf l is recorded. The recorded HB values are used to compute the adaptive tie, which uses the mean of the values for each leaf l instead of a fixed user-defined value as in VFDT. Using all the prediction statistics gathered in the testing phase for implementing prepruning, Pseudocode 3 presents the pseudocode of the training phase used by OVFDT for building an upright tree.

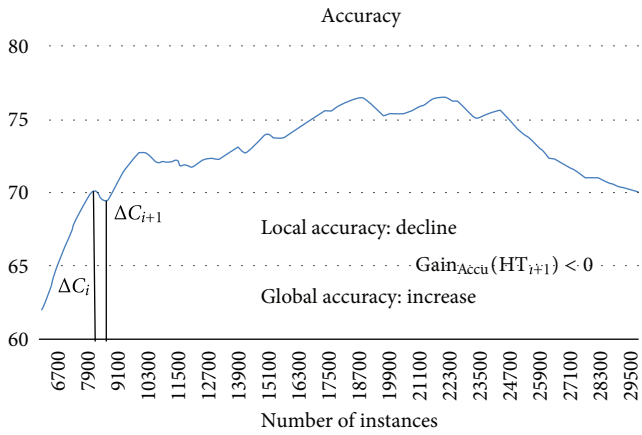


Figure 2: Example of incremental pruning.

3. Experiments

3.1. Bias Generator. For our experiments, we adopted and customized massive online analysis (MOA), one of the most popular data stream-mining toolkits, by including the aforementioned modifications into the OVFDT algorithm. However, the latest version of the MOA simulation environment is not able to simulate a biased data stream with an imbalanced class. A bias generator was therefore written in JAVA code and integrated into MOA for the purpose of evaluating the performance of stream-mining algorithms under imbalanced class data. Using either a simple command-line console or a graphic user, of which an example is shown in Figure 3, the generator injected biased instances from a specific imbalanced class into a given ARFF file. The input parameters are as follows:

(i) *biased class index* (BCI): the class index that the bias-added instances belong to,

(ii) *bias change from class index* (CCI): the class index that the bias instances will replace,

(iii) *change reduction percentage* (CP): the proportion of instances that will change to biased instances.

After the generator configuration, the instances with CCI class are replaced by BCI instances according to the CP setting. For example, in the snapshot below BCI = 5, CCI = 4, and CP = 80%; this means that 80% of CCI instances are replaced by BCI instances (the original settings are BCI = 10% and CCI = 10%, after BCI = 18% and CCI = 2%).

3.2. Experiment Datasets and Visualization. Six datasets were used to test the performance of OVFDT + FL versus ordinary VFDT. The datasets included those generated by the biased simulator and naturally imbalanced real-life data downloaded from the UCI machine-learning archive (http://www.ics.uci.edu/~mlearn). Table 1 describes these experimental datasets in detail, and Figure 4 provides a group of class distribution visualizations.

The following charts visualize the bias-included datasets with the imbalanced class. The pie charts on the left show the class distribution in the full experimental datasets and the charts on the right show the class distribution being progressively updated as new data streams arrive. For example, from Table 1 we see that the biased classes in the LED24 dataset are Class 2 (18%) and Class 4 (18%). There are 80% more data samples for these two classes than for the other classes. Originally the data distributions over all classes were equal. The charts representing the other datasets show at least one class, which has a larger percentage of data distribution than others.

3.3. Experiment Results Comparing VFDT and OVFDT. VFDT is deemed to be a suitable candidate for real-time classification in wireless sensor networks, because of its incremental learning nature based on a test-and-train approach. In this paper, we extend the design of VFDT to OVFDT, which has superior mechanisms for dealing with imbalanced data classes. This following comparison is between VFDT and OVFDT, which use the same types of functional tree leaf in the imbalanced datasets. The goal is to observe the comparative impact of the imbalanced classes on VFDT and OVFDT. For VFDT, the fixed tie breaking threshold (range from 0 to 1) is an important predefined parameter τ, which controls the node-splitting speed. In the experiment, τ was set at different values from 0.1 to 1.0 to test several different trails of VFDT, as a priori information for τ values is unavailable until the model is actually put to the test. The number of correctly classified instances measures the accuracy over the total number of arrived instances.

The results show, on the one hand, that $\text{OVFDT}_{\text{Adaptive}}$ has better performance results than any other method, for imbalanced data streams. The highlighted areas in Figure 5 show that OVFDT consistently outperformed VFDT. OVFDT_{MC} had lower accuracy than other functional tree leaf strategies in OVFDT. The advantage of the functional tree leaf approach is more apparent in the analysis of imbalanced data streams that have a significantly large bias in class distribution. This means that the modification at the testing phase is substantially effective, even when processing highly imbalanced data classes. On the other hand, the advantage of the other two modifications to the training phase, prepruning and dynamic node splitting, shows their usefulness in reducing the overfitting problem caused by imbalanced class data streams.

The radar chart in Figure 6 demonstrates that OVFDT results in a much smaller tree size than VFDT in all cases. A small tree size means lower runtime memory requirements, which makes it suitable for operating sensor node devices in WSNs. Tree size is measured by the number of leaves in a decision tree. Ideally there should be just enough leaves and corresponding branch paths to correctly classify the samples. Having too many leaves is a symptom of overfitting, which results in a decision tree that cannot make meaningful predictions and uses up memory space.

As these experimental results show, OVFDT with a functional tree leaf handles imbalanced data streams more effectively than VFDT. For this reason, VFDT will not be considered in the following experiments. Instead, we will

PROCEDURE doNodeSplittingEstimation(ΔC, S,X, $G(\cdot)$, δ)
(1) FOR each attribute $X_i \in X_l - \{X_\varnothing\}$ at the leaf l
(2) Compute $G_l(X_i)$
(3) Let X_a be the attribute with highest $G_l(\cdot)$ and X_b, with the 2nd highest $G_l(\cdot)$
(4) Compute HB with δ
(5) Let $\Delta G_l = G_l(X_a) - G_l(X_b)$
(6) END-FOR
(7) IF ($\Delta G_l >$ HB) or ($\Delta G \leq T_l^{\text{LOWER}}$ and $\Delta C_i < \Delta C_{i-1}$) or ($\Delta G \leq T_l^{\text{LOWER}}$ and $\Delta C_i < 0$) or ($T_l^{\text{LOWER}} < \Delta G \leq T_l^{\text{UPPER}}$ and $\Delta C_i < \Delta C_{i-1}$)
(8) Replace l by an internal node splits on X_a
(9) Update Adaptive tie T_l^{LOWER} and T_l^{UPPER}
(10) FOR each branch of splitting
(11) Add a new leaf l_m and let $X_m = X - \{X_a\}$
(12) Let $G(X_\varnothing)$ be $G(\cdot)$ obtained by predicting the class in S, according to $\mathcal{F}$ at l_m
(13) FOR each class y_k and each value x_{ij} of each attribute
(14) $X_i \in X_m - \{X_\varnothing\}$ and reset OCD: $n_{ijk}(l) = 0$
(15) END-FOR
(16) END-FOR
(17) END-IF
(18) Return updated HT

Pseudocode 3: The pseudocode of OVFDT model training.

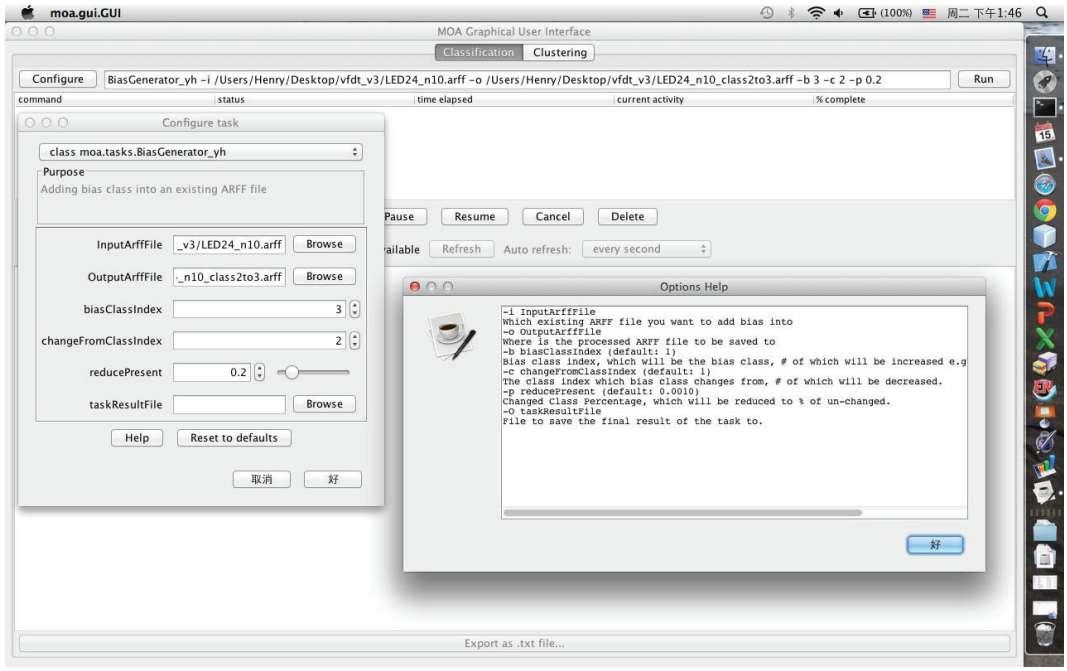


Figure 3: Snapshot of the Bias generator for generating data with imbalanced class, on MOA platform.

Table 1: Datasets with imbalanced class used in the experiment.

Name	Type	Source	Nom. attr. no.	Num. attr. no.	CLS no.	Bias CLS	Inst no.
LED24	Nominal	Synthetic	24	0	10	2, 4	10^6
Connect-4	Nominal	UCI	42	0	7	"Draw"	67,557
Waveform 21	Numeric	Synthetic	0	21	3	1	10^6
Radial bias function (RBF)	Numeric	Synthetic	0	50	10	1, 3, 5, 8, 10	10^6
Random tree (RT)	Mixed	Synthetic	50	50	10	4, 5, 10	10^6
COVTYPE	Mixed	UCI	42	12	7	2, 7	581,012

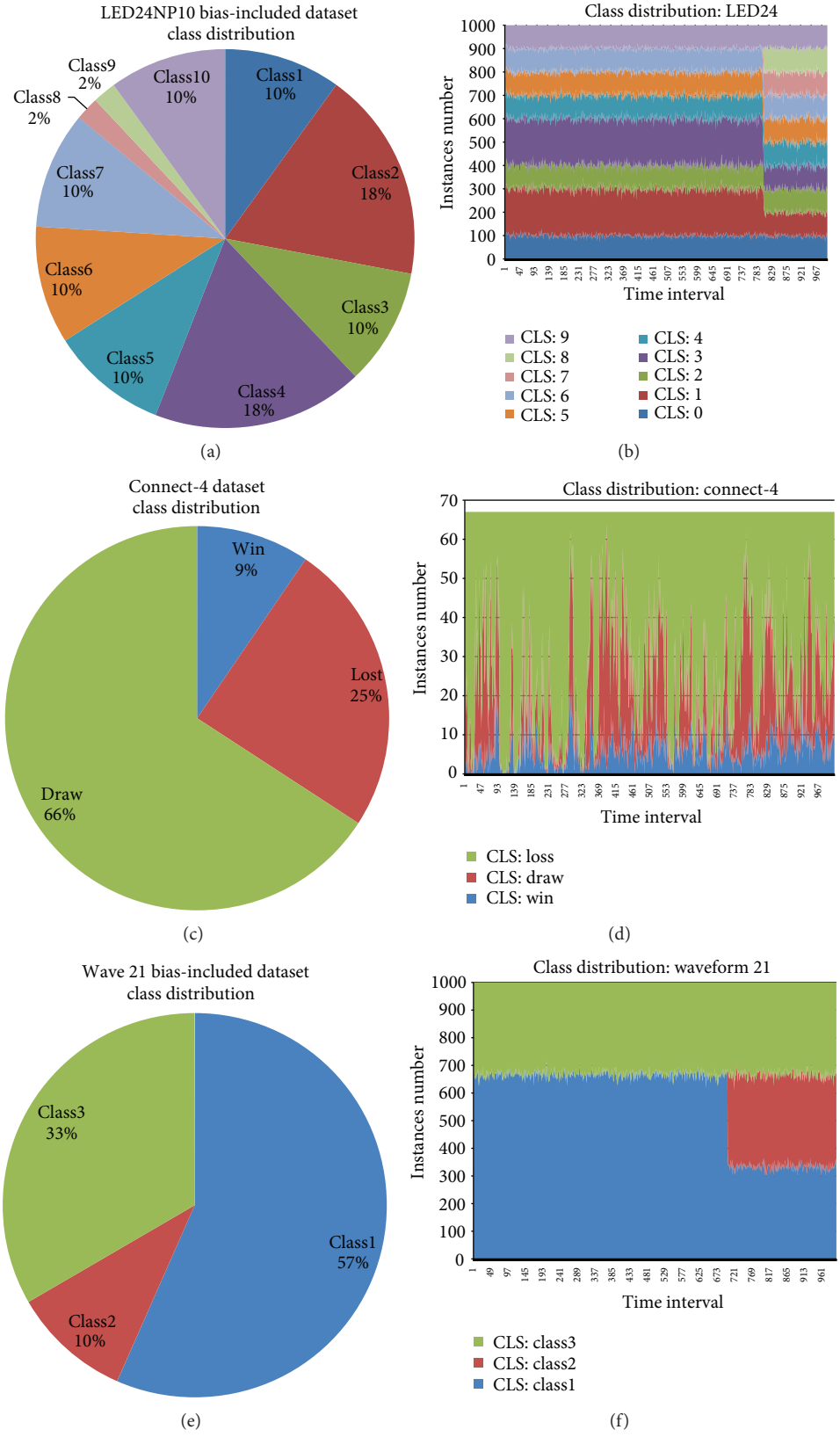


FIGURE 4: Continued.

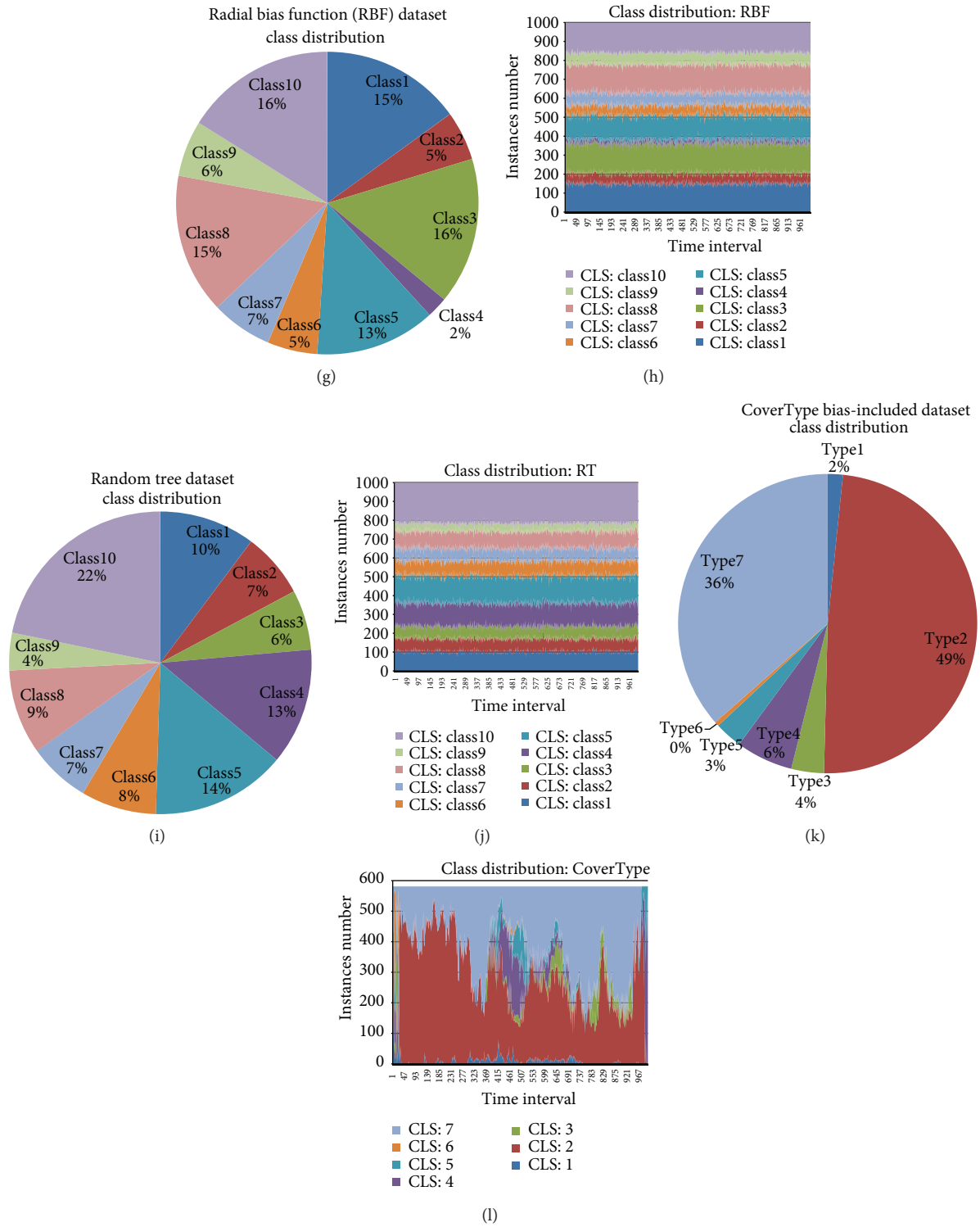


Figure 4: A collection of visualizations of the datasets that have different degrees of imbalanced class distribution.

analyze in detail the experimental results of OVFDT using different types of functional tree leaves.

3.4. Experiment Results Comparing OVFDT Functional Tree Leaf Accuracy. Comparing the classification accuracy of four different types of functional tree leaves, we find that OVFDT_{MC} always obtains the lowest accuracy and $\text{OVFDT}_{\text{Adaptive}}$ has consistently better accuracy than the other methods. In addition, $\text{OVFDT}_{\text{WNB}}$ is better than OVFDT_{NB} in experiments that weight the probabilities of each attribute occurrence (see Figure 8).

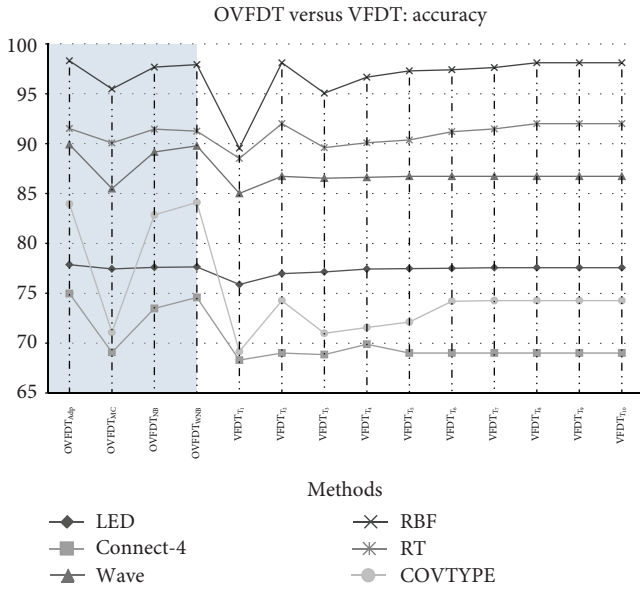


Figure 5: Accuracy of the classification experiments by VFDT and OVFDT with datasets of imbalanced data class.

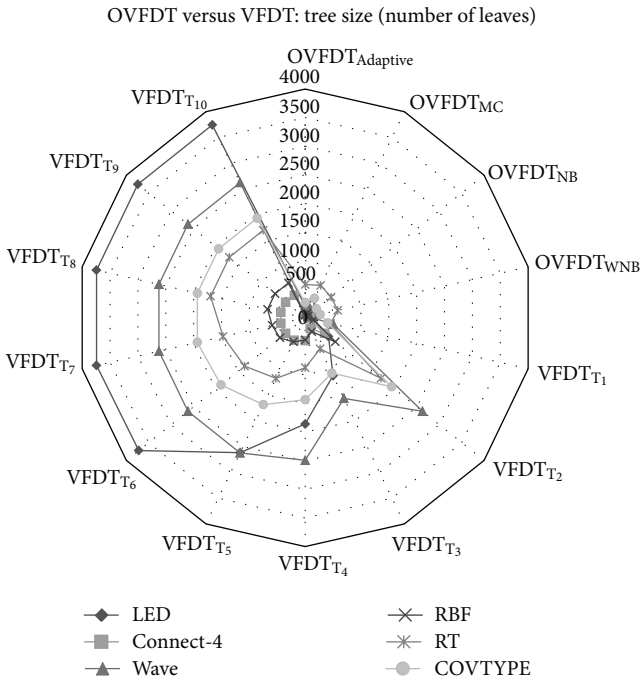


Figure 6: Tree size of the classification experiments by VFDT and OVFDT with datasets of imbalanced data class.

Another good performance benchmark is the receiver operating characteristic (ROC), which is a standard method for analyzing and comparing classifiers when the costs of misclassification are unknown. In a stream-mining scenario, it is not possible to know the misclassification costs, because the mining process is incremental over running data streams, and does not analyze a full dataset. The ROC provides a convenient graphical display of the tradeoff between the true

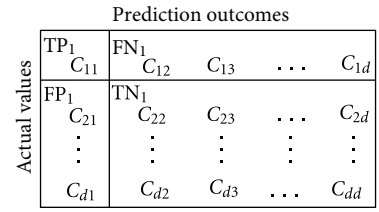


Figure 7

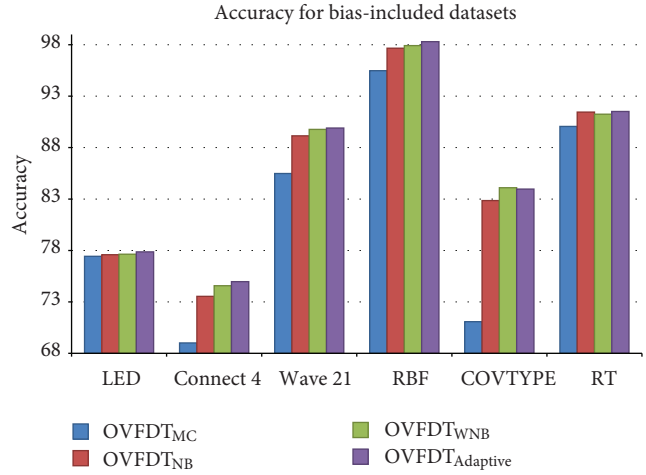


Figure 8: Tree size of the classification experiments by different FL types of OVFDT with datasets of imbalanced data class.

and false positive classification rates for two class problems [11]. In the decision tree classification, however, there are more than two classes. Therefore, we extend the standard ROC model to a multiclass ROC analysis to evaluate the tree learning algorithm's performance.

Suppose that there is a D-class classification system, with d-dimensional classes that need to be classified by the tree learning algorithm. A $d \times d$-dimensional confusion matrix or contingency table C, which summarizes the results of the classifications, presents the true positives and false positives for the multi-class analysis. Each entry C_{ii} of the matrix C gives the number of examples, whose true class was A_i, that were actually assigned to A_i, where $1 \leq i \leq d$. Each entry C_{ij} of the matrix C gives the number of examples, whose true class was A_i, that were actually assigned to A_j, where $i \neq j$ and $1 \leq i, j \leq d$:

$$C = \begin{bmatrix} C_{11} & \cdots & C_{1d} \\ \vdots & \ddots & \vdots \\ C_{d1} & \cdots & C_{dd} \end{bmatrix}. \quad (9)$$

To use two-class ROC statistics, each class i to d in the multi-class ROC is assigned a negative or positive value. Samples with class i are positive; otherwise, negative. True positives (TP) are examples correctly labeled as positives. False positives (FP) refer to negative examples incorrectly labeled as positive. True negatives (TN) are negatives correctly labeled as negative. Finally, false negatives (FN) refer to positive examples incorrectly labeled as negative. Each

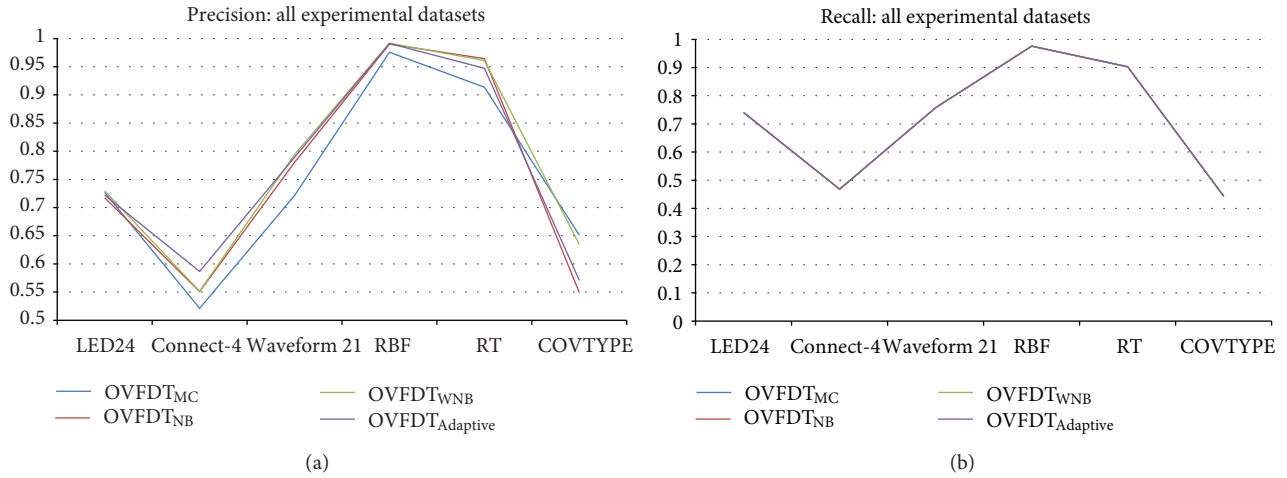


Figure 9: Precision and Recall values of the classification experiments by different FL types of OVFDT with datasets of imbalanced data class.

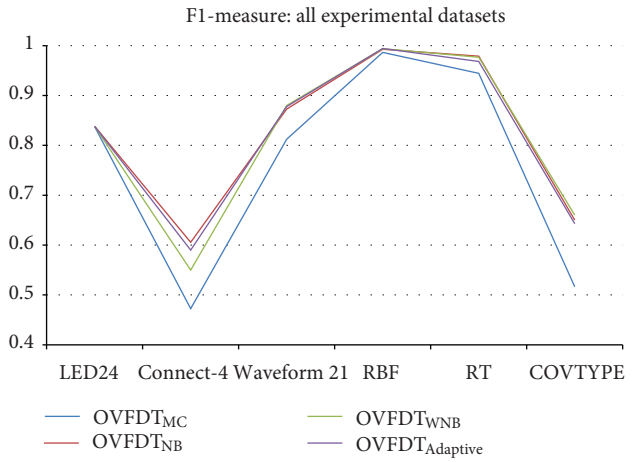


Figure 10: F1-measure of the classification experiments by different FL types of OVFDT with datasets of imbalanced data class.

class i can be converted into a two-class problem, with the corresponding values of *True Positive* (10), *False Positive* (11), *False Negative* (12), and *True Negative* (13) (see Figure 7):

$$\mathrm{TP}_i = C_{ii}\,, \tag{10}$$

$$\mathrm{FP}_i = \left(\sum_{j=1}^{d} C_{ji}\right) - C_{ii}, \tag{11}$$

$$\mathrm{FN}_i = \left(\sum_{j=1}^{d} C_{ij}\right) - C_{ii}, \tag{12}$$

$$\mathrm{TN}_i = \left(\sum_{i=1}^{d}\sum_{j=1}^{d} C_{ij}\right) - \mathrm{TP}_i - \mathrm{FP}_i - \mathrm{FN}_i. \tag{13}$$

Precision-Recall is a well-known method of analyzing ROC. In pattern recognition, precision is the fraction of retrieved instances that are relevant, while recall is the fraction of relevant instances that are retrieved. The values of precision and recall range from 0 to 1. A precision score of 1 for a class i means that every item labeled as belonging to class i does indeed belong to class i. A recall score of 1 means that every item from class i was labeled as belonging to class i. Precision-Recall scores are not analyzed in isolation. F_β-measure [12] is a weighted harmonic mean of the Precision-Recall measure. The F1-measure evenly weights precision and recall scores. The best value for the F1-measure is 1 and the worst score is 0. In addition, the true positive rate (TPR) and the false positive rate (FPR) are common benchmarks in ROC analysis:

$$\begin{aligned}
\mathrm{Precision}_i &= \frac{\mathrm{TP}_i}{(\mathrm{TP}_i + \mathrm{FP}_i)} = \frac{C_{ii}}{\left(\sum_{j=1}^{d} C_{ji}\right)},\\
\mathrm{Recall}_i &= \frac{\mathrm{TP}_i}{(\mathrm{TP}_i + \mathrm{FN}_i)} = \frac{C_{ii}}{\sum_{j=1}^{d} C_{ij}},\\
F_1\ \mathrm{Measure}_i &= \frac{2\mathrm{TP}_i}{\mathrm{TP}_i + \mathrm{FN}_i + \mathrm{TP}_i + \mathrm{FP}_i}\\
&= \frac{2C_{ii}}{\sum_{j=1}^{d} C_{ij} + \sum_{j=1}^{d} C_{ji}}.
\end{aligned} \tag{14}$$

We analyze the Precision-Recall for each class for all the imbalanced class datasets. Due to limited space, the detailed charts are given in the Appendix. The average Precision-Recall values are described in Figures 9, and 10.

These charts illustrate that the average precision of OVFDT$_{MC}$ is worse than those of the other methods. OVFDT$_{Adaptive}$ obtains the highest precision in homogenous (nominal only and numeric only) datasets. All methods have the same average values of recall (the lines appear to overlap). We then apply the F1-measure to evaluate the experiment result. As the chart below shows, the value ranges from 0 to 1; OVFDT$_{MC}$ again has the lowest F1-measure value. However, because the datasets

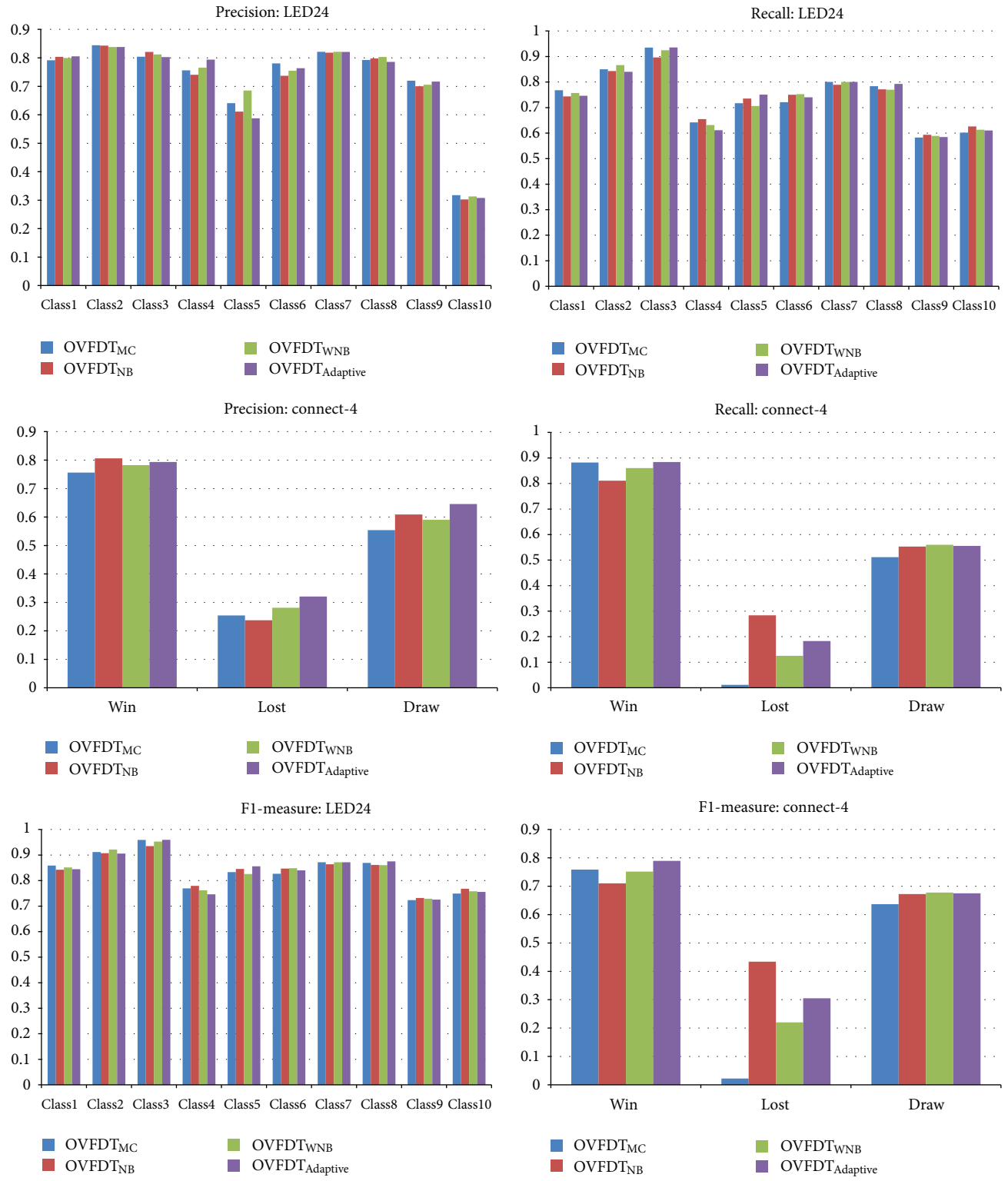


Figure 11

contained biased instances in imbalance classes, the average Precision-Recall analysis is not sufficient to determine accuracy. We must consider the Precision-Recall for the distributed class in every different data stream. From the above experiments, we observe that $OVFDT_{Adaptive}$ always achieves higher precision, recall, and F1-measure values than $OVFDT_{MC}$, but this was not the case for $OVFDT_{NB}$ and $OVFDT_{WNB}$.

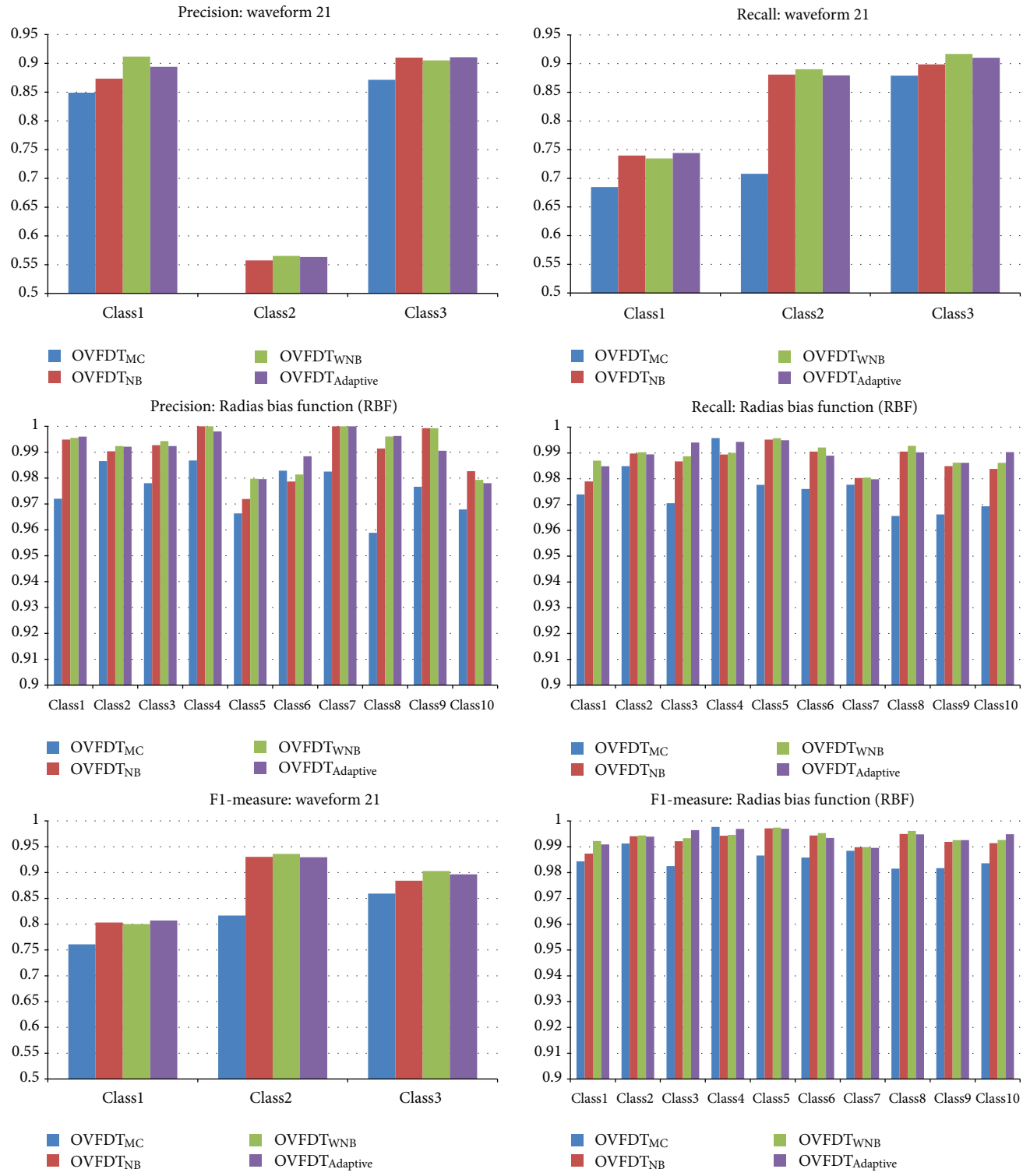


Figure 12

4. Conclusion

Imbalanced data classification is a challenging problem that generally refers to a learning model created for a dataset that has far more samples in one class than in the others. In an ubiquitous environment such as a wireless sensor network, it is not uncommon for the data of interest to fall into a small minority class. Previous researchers have tackled this problem by using techniques that inevitably create additional computation overheads. These techniques usually include

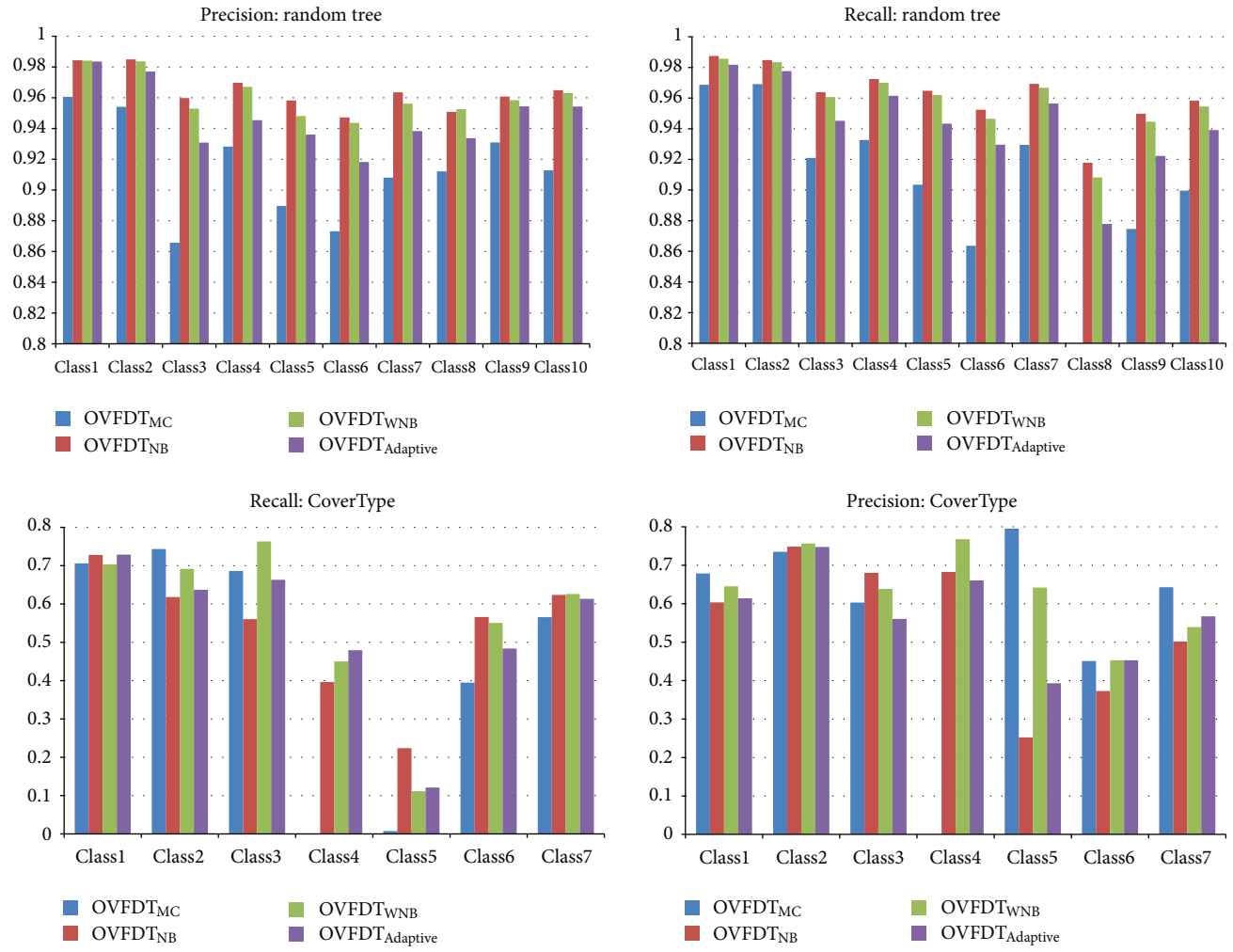


Figure 13

resampling the observations from a bounded archive so as to balance the imbalance. Others may resort to postpruning the decision tree and redistributing the classification costs in a backward-learning process. All of these proposed techniques worked well in traditional data mining but might not suit a real-time stream-mining scenario, where all the data arrive in a single pass; at a sensor sink it is neither practical nor feasible to archive a stationary set of data, let alone to resample.

In this paper, a novel solution is introduced at the algorithmic level, which is based on a popular stream-mining algorithm called the Very Fast Decision Tree (VFDT). Three modifications are proposed for VFDT as a means to reduce the effect of imbalanced class data. The modifications are implemented at the training phase prior to expanding the decision tree and at the testing phase, where prediction accuracy is fine-tuned by weighting the leaves of the decision trees according to the probabilities of the arriving data. The overall solution is called the Optimized VFDT with Functional Tree Leaf (OVFDT + FL). The mechanism of Function Tree Leaf is implemented by using weighted naïve Bayes predictors, installed at the decision tree leaves of the OVFDT. Specifically, perturbed datasets that include "biased" class distribution are used for experiments for illustrating the efficacy of the new algorithm. OVFDT + FL is shown to outperform VFDT in a series of experiments where datasets are deliberately biased by a custom-made data generator software program. In particular, two variants of FL called adaptive and weighted naïve Bayes performed consistently better than other techniques. OVFDT succeeded in minimizing the impacts of imbalanced class data, while maintaining high accuracy and a compact decision tree size. This contrasts with the known over-fitting problems of poor accuracy and huge tree size usually caused by imbalanced class data. The OVFDT + FL is validated as a good classification model for wireless sensor networks.

Appendix

A. Precision-Recall Charts for Each Dataset

A.1. Homogenous Data: Nominal Only. See Figure 11.

A.2. Homogenous Data: Numeric Only. See Figure 12.

A.3. Both Nominal and Numeric Attributes. See Figure 13.

Acknowledgments

The authors are thankful for the financial support from the research Grant "Real-time Data Stream Mining," Grant no. RG070/09-10S/FCC/FST, offered by the University of Macau, FST, and RDAO.

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, no. 4, pp. 393–422, 2002.

[2] M. Di and E. M. Joo, "A survey of machine learning in wireless sensor netoworks—from networking and application perspectives," in *Proceedings of the 6th International Conference on Information, Communications and Signal Processing (ICICS '07)*, pp. 1–5, Singapore, December 2007.

[3] Y. L. Borgne and G. Bontempi, "Round robin cycle for predictions in wireless sensor networks," in *Proceedings of the 2nd International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, pp. 253–258, December 2005.

[4] J. Zhang, E. Bloedorn, L. Rosen, and D. Venese, "Learning rules from highly unbalanced data sets," in *Proceedings of the 4th IEEE International Conference on Data Mining (ICDM '04)*, pp. 571–574, November 2004.

[5] J. Yu, Y. Qi, G. Wang, Q. Guo, and X. Gu, "An energy-aware distributed unequal clustering protocol for wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 2011, Article ID 202145, 8 pages, 2011.

[6] H. M. Nguyen, E. W. Cooper, and K. Kamei, "Borderline over-sampling for imbalanced data classification," *International Journal of Knowledge Engineering and Soft Data Paradigms*, vol. 3, no. 1, pp. 4–21, 2011.

[7] J. Wang, M. Xu, H. Wang, and J. Zhang, "Classification of imbalanced data by using the SMOTE algorithm and locally linear embedding," in *Proceedings of the 8th International Conference on Signal Processing (ICSP '06)*, pp. 16–20, November 2006.

[8] Y. Zhai, N. Ma, B. An, and D. Ruan, "An effective over-sampling method for imbalanced data sets classification," *Chinese Journal of Electronics*, vol. 20, no. 3, pp. 489–494, 2011.

[9] P. Domingos, "MetaCost: a general method for making classifiers cost-sensitive," in *Proceedings of the 5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 155–164, San Diego, Calif, USA, 1999.

[10] M. Kubat, R. C. Holte, and S. Matwin, "Machine learning for the detection of oil spills in satellite radar images," *Machine Learning*, vol. 30, no. 2-3, pp. 195–215, 1998.

[11] Y. Hang, S. Fong, G. Sun, and R. Wong, "A very fast decision tree algorithm for real-time data mining of imperfect data streams in a distributed wireless sensor network," *International Journal of Distributed Sensor Networks*, vol. 2013, Article ID 863545, 2013.

[12] J. A. O. Gama, R. Rocha, and P. Medas, "Accurate decision trees for mining high-speed data streams," in *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '03)*, pp. 523–528, New York, NY, USA, August 2003.

[13] Y. Hang and S. Fong, "OVFDT with functional tree leaf—majority class, naive bayes and adaptive hybrid integrations," in *Proceedings of the 3rd International Conference on Data Mining and Intelligent Information Technology Applications (ICMIA '11)*, IEEE Press, Macau, China, October 2011.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 532572, 5 pages
doi:10.1155/2012/532572

Research Article

Multiple Interface Parallel Approach of Bioinspired Routing Protocol for Mobile Ad Hoc Networks

L. J. García Villalba,[1] D. Rupérez Cañas,[1] A. L. Sandoval Orozco,[1] and T.-H. Kim[2]

[1] *Grupo de Análisis, Seguridad y Sistemas (GASS), Departamento de Ingeniería del Software e Inteligencia Artificial (DISIA), Facultad de Informática, Universidad Complutense de Madrid (UCM), Despacho 431, Calle Profesor José García Santesmases s/n, Ciudad Universitaria, 28040 Madrid, Spain*
[2] *School of Information Science, GVSA and UTAS, 20 Virginia Court, Sandy Bay, Hobart, TAS 7001, Australia*

Correspondence should be addressed to L. J. García Villalba, javiergv@fdi.ucm.es

Received 30 September 2012; Accepted 6 October 2012

Academic Editor: Sabah Mohammed

The design of routing protocols for mobile ad hoc networks (MANETs) is a complex task given the dynamic nature of such networks. Particular types of routing protocols are known as bioinspired. This work presents a parallelization of AntOR-DNR, a bioinspired routing protocol for mobile ad hoc networks based on the Ant Colony Optimization (ACO) algorithm. This new protocol, called PantOR-MI, uses, as well as PAntOR, the thread programming based on shared memory. This new parallelization is applied in route discovery phases, route local repair process, and link failure notification. The simulation results indicate that PAntOR and PAntOR-MI improve performances of AntOR, whilst it is also noticed that PAntOR-MI is the most suitable for highly dynamic environments.

1. Introduction

A mobile ad hoc network (MANET) [1] is a collection of mobile devices, which form a network of communication without predefined infrastructure. This fact determines the design of routing protocols for this type of network to suppose an arduous task. Particular types of routing protocols are called bioinspired [2], taking into account the behavior of some animals (insects, etc.) to obtain their food. Related to theses, the algorithms based on Ant Colony Optimization (ACO) [3], are particularly relevant. A representative protocol of so-called bioinspired is AntOR [4], adaptive and multihop routing protocol based on AntHocNet [5, 6]. The specification of this protocol includes two versions: disjoint-link routes (AntOR-DLR) and disjoint-node routes (AntOR-DNR). A parallel approximation of AntOR-DNR is PAntOR [7, 8] improving performances of AntOR-DNR through thread programming based on shared memory in the phases of routing information setup, route local repair, and link failure notification. This paper presents a new parallel approximation of AntOR-DNR, called PAntOR-Multiple Interface (PAntOR-MI) which, as its name suggests, differs essentially from PAntOR in the use of multiple interfaces. This paper consists of 6 sections, with this introduction being the first of them. The rest of the paper is structured as follows: Section 2 briefly discusses the most representative work on parallel techniques for bioinspired protocols based on the behavior of ants. Section 3 briefly comments on AntOR-DNR and PAntOR, also showing a comparison between both. Section 4 presents PAntOR-MI, with emphasis on differences from its predecessor, PAntOR. Section 5 shows a comparative study between AntOR-DNR, PAntOR, and PAntOR-MI. Finally, the conclusions are exposed in Section 6.

2. Related Works

In this section we present the most representative parallelization techniques that make ACO more efficient. First of which, introduced by [9], explains a method that can solve difficult combinatorial optimization problems. Stützle [10] applies an approximation master/slave to parallelize several different

search methods of ACO solutions. Reference [11] shows a hybrid system of parallelization which consists of evaluating the performance of communication multithreading Message Passing Interface (MPI). This approach applies MPI between nodes and multithreading within nodes. Finally, [7, 8] presents PAntOR, an approximation parallel based on programming by threads, which constitutes the starting point of the present paper. The main idea of this protocol is to replace each broadcast message by a message managed by a thread that is addressed to each one-hop neighbour, that is, launches a thread by each node in the neighbour table. This is done in protocol phases: routing information setup, route local repair, and link failure notification.

3. AntOR versus PAntOR

AntOR [4] has the following characteristics which are different from AntHocNet [5, 6]:

(i) Disjoint-link and Disjoint-node protocol,

(ii) separation between the pheromone values in the diffusion process, and

(iii) use distance metric in route proactive exploration.

PAntOR [7, 8] paralyze Disjoint-node version (AntOR-DNR) that consists of nodes from routes which are not shared. We have chosen this version because routes are more difficult to get and maintain. To understand how P-AntOR works, it is necessary to use three concepts well.

(a) *Process*. It is a program running which is managed by the Operating System.

(b) *Thread*. It consists of the basic unit of execution, so any program that executes has at least a thread.

(c) *POSIX Thread*. It is a Standard based in thread API for C/C++.

We use POSIX Thread because it allows a new concurrent process flow to expand, which is the most efficient multicore systems, generating a flow of processes that can be scheduled to run on another processor, thus speed through distributed processing is achieved. Programming with threads carries less overhead than expanding a new process, because the system does not initialize a new environment and virtual memory space for that process.

This version tries to replace each broadcast message by a message managed by a thread that is addressed to each one-hop neighbour.

This parallel technique, which launches a thread by each node in the neighbour table, is used in the following phases of AntOR.

3.1. Routing Information Setup. Figure 1 shows a flow chart representing the parallelism in the route discovery process. When a source node is ready to send data to the destination node, it activates the route discovery process. This process is parallelized using threads, so that it launches a Reactive Forward Ant (RFA) reactive to the one-hop neighbours

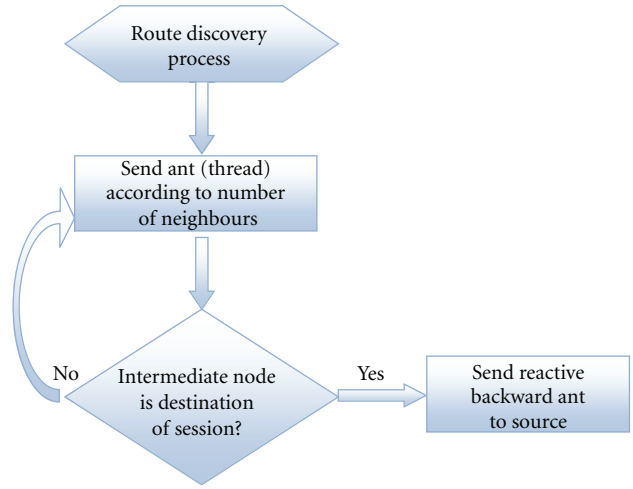


Figure 1: Parallelization of route discovery phase.

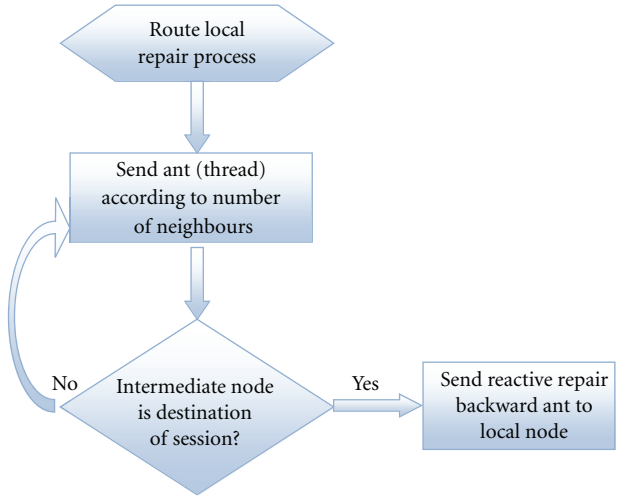


Figure 2: Parallelization of local route repair phase.

through an independent thread. When an intermediate node receives this ant repeats the process, but whether it is a destination node, this node sends its corresponding Reactive Backward Ant (RBA).

3.2. Route Local Repair. The operation is similar to route discovery, unless it is done locally, as shown in Figure 2.

3.3. Link Failure Notification. This process updates the routing table to link failures. It is required to be taken promptly because of its importance. The nodes send ants through independent threads until an intermediate node has some alternative route to the destination after updating the routing table, as shown Figure 3.

3.4. Results. Next we show a comparison between PAntOR and AntOR. To this end, the metrics used in the simulations were as follows.

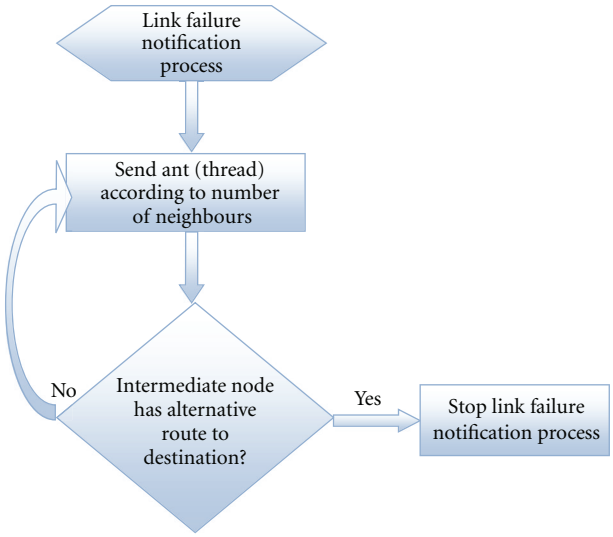


Figure 3: Parallelization of link failure notification phase.

(i) *Throughput*: consists of volume of work or information flowing through a system. It is calculated by dividing the total number of bits delivered to the destination by the packet delivery time.

(ii) *Delivered Data Packet Ratio*: relationship between number of packets sent and the number of packets delivered successfully.

(iii) *Overhead in number of packets*: relationship between the total numbers of transmitted control packets by the nodes of network and the number of delivered data packets to their destinations.

Experiments with the Network Simulator NS-3 have been realized. Simulations parameters are as follows: we have used 100 nodes configured according to the Standard IEEE 802.11b, moving in a random scenario of dimensions 1200 m × 1200 m according to the mobility pattern *Random WayPoint* (RWP). The application of data traffic is *Constant Bit Rate* (CBR) with a rate of packet sending of 2048 bps (4 packets of 64 bytes per second). We apply 5 random data sessions, where mobility is variable from 0 m/s up to a maximum of 10 m/s. Pause time is kept constant to a value of 30 s. Total simulation time is 120 s. In Figure 4 we can observe how the Throughput is better at PAntOR at AntOR.

In Figure 5 see how overhead in number of packets is better in the parallel version than in the original, because the creation of routes is faster.

Figure 6 behaves similar to the representation in Figure 4, but using another scale. It notes that the performance of the packet delivery remains reasonably good even at high speeds.

4. PAntOR-Multiinterface

This variant of PAntOR consists of having more than one interface, and to parallelize the sending of broadcast messages by interface through threads. Each interface is managed by a thread. The main idea of this parallelization

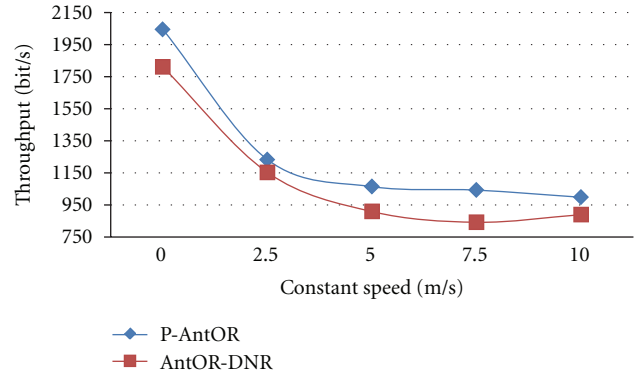


Figure 4: Constant speed against Throughput.

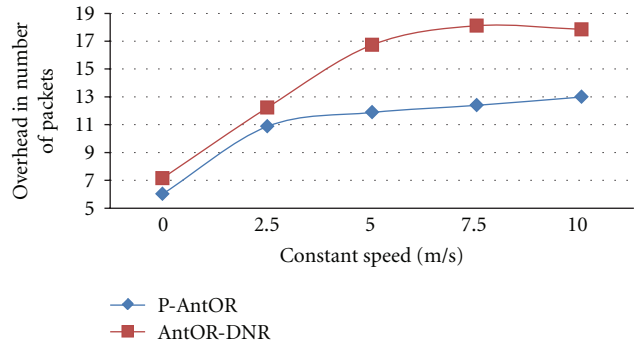


Figure 5: Constant speed against overhead in number of packets.

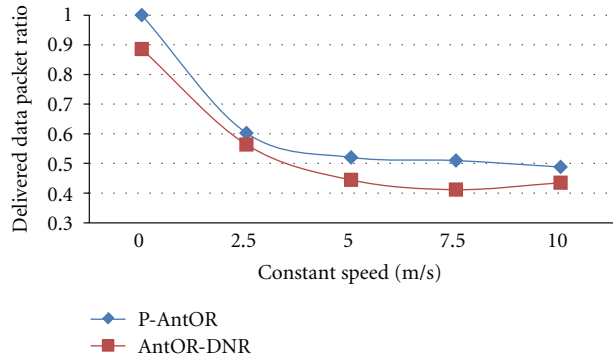


Figure 6: Constant speed against Delivered Data Packet Ratio.

is to be applied to systems with several multiinterfaces, which launches an ant broadcast mode in an independent thread for each interface that provides the node. The main difference with PAntOR is that PAntOR-MI uses more than one interface, parallelizing each interface by means of a thread.

To understand this variant we provide Algorithm 1.

It can be seen that while running the routing information setup, a reactive message is launched in a broadcast way for each interface that have the node, and such an interface via a thread is managed.

```
while Routing Information Setup do
    for Cont = 1 to Max_Interfaces do
        Send Broadcast Message by Thread (Cont);
    end
end
```

Algorithm 1: Parallelization core in PAntOR-MI.

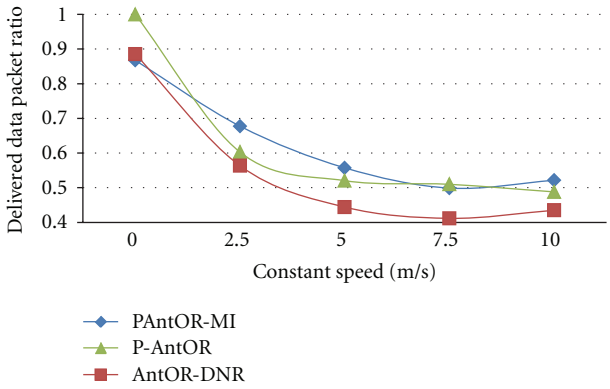


Figure 7: Constant Speed against Delivered Data Packet Ratio.

5. PAntOR-MI versus AntOR and PAntOR

The performance metric analyzed in this comparison is Delivered Data Packet Ratio, which consists of the relationship between the number of packets sent and the number of packets delivered correctly. To perform this comparison Network Simulator NS-3 has been used with the following parameters: a random scenario with dimensions of 1200 m $\times$ 1200 m has been designed, where 100 nodes configured according to the Standard IEEE 802.11b; they move according to the mobility pattern *Random WayPoint* (RWP). The application of data traffic is based on *Constant Bit Rate* (CBR) with a packet sending rate of 2048 bps (4 packets of 64 bytes per second). We apply 5 random data sessions, where mobility is variable from 0 m/s up to a maximum of 10 m/s. Pause time is kept constant to a value of 30 s. Total simulation time is 120 s. In this comparison PAntOR-MI use nodes with two interfaces.

In Figure 7 we appreciate how the delivered packet ratio is better in these two parallel versions than in AntOR. Also we see how PAntOR-MI improves to AntOR and PAntOR. Also a greater tolerance to the mobility of the nodes is observed, behaving better in the more dynamic scenarios.

6. Conclusions

This work has presented a new bioinspired routing protocol for mobile ad hoc networks obtained thanks to new parallelization techniques of a base protocol called AntOR which has two versions, the so-called Disjoint-link (AntOR-DLR) and Disjoint-node (AntOR-DNR). The new parallel approach (PAntOR-MI) used the disjoint-node version of AntOR (AntOR-DNR) as the main protocol, as well as the existing (PAntOR). The parallelization technique employed is a large-grained approach, in which a multicore machine in a shared memory system has been used. The novelty of PAntOR-MI is that we have more than one interface, and we parallelize the sending of broadcast messages by interface using threads. Each interface is managed by a thread. The obtained simulation results indicate that PAntOR-MI improves performances of PAntOR, with further observations showing that this improvement is most evident in most dynamic environments.

Acknowledgment

This work was supported by the Agencia Española de Cooperación Internacional para el Desarrollo (AECID, Spain) through Acción Integrada MAEC-AECID Mediterráneo A1/037528/11.

References

[1] M. Abolhasan, T. Wysocki, and E. Dutkiewicz, "A review of routing protocols for mobile Ad Hoc networks," *Ad Hoc Networks*, vol. 2, no. 1, pp. 1–22, 2004.

[2] V. Jha, K. Khetarpal, and M. Sharma, "A survey of nature inspired routing algorithms for MANETs," in *Proceedings of the 3rd International Conference on Electronics Computer Technology (ICECT '11)*, pp. 16–24, Kanyakumari, India, April 2011.

[3] M. Dorigo and T. Stützle, *Ant Colony Optimization*, The MIT Press, Bradford Company Scituate, Mass, USA, 2004.

[4] L. J. G. Villalba, D. R. Cañas, and A. L. S. Orozco, "Bio-inspired routing protocol for mobile Ad Hoc networks," *IET Communications*, vol. 4, no. 18, pp. 2187–2195, 2010.

[5] F. Ducatelle, *Adaptive routing in Ad Hoc wireless multi-hop networks [Ph.D. thesis]*, Università della Svizzera Italiana, Istituto Dalle Molle di Studi Sull'Intelligenza Artificiale, 2007.

[6] G. Di Caro, F. Ducatelle, and L. M. Gambardella, "AntHocNet: an adaptive nature-inspired algorithm for routing in mobile Ad Hoc networks," *European Transactions on Telecommunications*, vol. 16, no. 5, pp. 443–455, 2005.

[7] L. J. García Villalba, D. Rupérez Cañas, and A. L. Sandoval Orozco, "Parallel approach of a bioinspired routing protocol for MANETs," *International Journal of Ad Hoc and Ubiquitous Computing*. In press.

[8] D. Rupérez Cañas, A. L. Sandoval Orozco, L. J. García Villalba, and T.-H. Kim, "Comparing AntOR-disjoint node routing protocol with its parallel extension," in *Proceedings of the International Conference on Multimedia, Computer Graphics and Broadcasting (MulGraB '11)*, pp. 305–309, 2011.

[9] B. Bullnheimer, G. Kostis, and Strauss, "Parallelization strategies for the ant systems," in *High Performance Algorithms and*

Software in NonLinear Optimization Series: Applied Optimization, vol. 24, 1998.

[10] T. Stützle, "Parallelization strategies for ant colony optimization," in *Proceedings of the Parallel Problem Solving from Nature*, vol. 1498, 1998.

[11] R. Thakur and W. Gropp, "Test suite for evaluating performance of multithreaded MPI communication," *Parallel Computing*, vol. 35, no. 12, pp. 608–617, 2009.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 818054, 6 pages
doi:10.1155/2012/818054

Review Article

Security Issues in Mobile Ad Hoc Networks

A. L. Sandoval Orozco,[1] J. García Matesanz,[2] L. J. García Villalba,[1] J. D. Márquez Díaz,[3] and T.-H. Kim[4]

[1] *Grupo de Análisis, Seguridad y Sistemas (GASS), Departamento de Ingeniería del Software e Inteligencia Artificial (DISIA), Facultad de Informática, Universidad Complutense de Madrid (UCM), Despacho 431, Calle Profesor José García Santesmases s/n, Ciudad Universitaria, 28040 Madrid, Spain*
[2] *Grupo de Análisis, Seguridad y Sistemas (GASS), Sección Departamental de Sistemas Informáticos y Computación, Lenguajes y Sistemas Informáticos y Ciencias de la Computación e Inteligencia Artificial-Facultad de Ciencias Matemáticas, Universidad Complutense de Madrid (UCM), Despacho 310-F, Plaza de Ciencias, 3, Ciudad Universitaria, 28040 Madrid, Spain*
[3] *Grupo de Redes de Computadores e Ingeniería de Software (GRECIS), Departamento de Ingeniería de Sistemas, Universidad del Norte, Km 5 Autopista a Puerto Colombia, Barranquilla, Colombia*
[4] *School of Information Science, GVSA and UTAS, 20 Virginia Court, Sandy Bay, Hobart, TAS 7001, Australia*

Correspondence should be addressed to T.-H. Kim, taihoonn@daum.net

Received 30 September 2012; Accepted 6 October 2012

Academic Editor: Sabah Mohammed

Ad hoc networks are built on the basis of a communication without infrastructure and major investigations have focused on the routing and autoconfiguration problems. However, there is a little progress in solving the secure autoconfiguration problems in mobile ad hoc networks (MANETs), which has led to the proliferation of threats given the vulnerabilities of MANETs. It is clear that ad hoc networks have no centralized mechanism for defense against threats, such as a firewall, an intrusion detection system, or a proxy. Therefore, it is necessary that the defense of interests of each of the ad hoc components is the responsibility of each member node. This paper shows the most common threats to ad hoc networks and reviews several proposals that attempt to minimize some of these threats, showing their protection ability and vulnerabilities in light of the threats that might arise.

1. Introduction

MANET technology is used to immediately provide secure access between multiple mobile nodes without the need for a preset communications infrastructure achieving a multihop architecture. These networks are identified by two basic principles: routing and autoconfiguration.

While there is already quite a lot of established work undertaken on routing [1–4] and consequently those related to secure routing [5–8], there is still a room for continuous improvements on those which are still under construction, notably those related to auto-configuration and in particular, those in connection with secure MANET auto-configuration. Thus, this paper shows the most important works carried out concerning the latter.

Insertion of a node to the MANET involves implementing initial configuration mechanisms [9, 10], such as assigning an available IP address before this node can participate actively in the network. There are three types of solutions to carry out this assignment: *stateful*, *stateless* or *hybrid*.

In *stateful* solutions, addresses are assigned by the network; therefore the network should maintain the status information of addresses that have been assigned and/or released.

In *stateless* solutions, the addresses are assigned by the same node that enters the MANET. This node should run a test for duplicate address detection (DAD) in order to determine the uniqueness of the assigned address.

The *hybrid* solutions combine aspects of both previous types of solutions to improve the scalability and reliability of auto-configuration mechanisms.

All proposals have advantages and disadvantages in terms of solving the following problems: uniqueness of addresses, network initialization, node departure, network partitioning and network merging. However, all lack a mechanism to ensure the authenticity of the address owner at the time in which the auto-configuration is carried out. As a result, a malicious node can spoof any node already set up to hijack its traffic, preventing other nodes from entering the network, sending messages with false addresses, causing denial of service by flooding the network with unsolicited messages from fake addresses,rejecting the possibility that other nodes can access the network, or causing the refusal to accept the insertion of a new node, when the auto-configuration mechanism requires that all nodes confirm the entry of a new member to the MANET.

Although studies over the authenticity of the nodes entering the MANET during auto-configuration have been minimal, the aim of this paper is to show how they have presented some solutions to this problem and show some of its shortcomings from the perspective of the characteristics to be evaluated for potential threats within the auto-configuration process.

This piece of work, including the introduction, is organized into four sections as follows. Section 2 shows an overview of possible threats that may occur within a MANET. In Section 3, the highlights of some proposed solutions to secure MANET auto-configuration are reviewed and analyzed. Finally, conclusions are presented in Section 4.

2. Threats in Autoconfiguration

In the processes applied during the execution of the mechanisms of auto-configuration, predictable and reliable behaviour from the nodes that compose the MANET is expected, as much from those which enter as from those already inside. However, this is not always the case, as malicious nodes can potentially be causing some damage, such as interference of messages, node impersonation, denial of service, spoofing, and eavesdropping among others.

In this paper we use the classification proposed by Wang et al. [11] and Buiati et al. [12] to specify the security threats.

(i) *Address Spoofing Threat*. A malicious node may deliberately choose an assigned or a free IP address for their attack. In the first case, the malicious node teases any configured node as its victim and hijacks its traffic, and in the second one, the node assigns the free IP address to itself to participate in the network, gathering important information necessary to execute active attacks, such as denial of service.

(ii) *Address Space Exhaustion Threat*. A malicious node can claim as many IP addresses as possible until exhausting the address space. This node may request the assignment of addresses to a ghost node (fake nodes). This way the malicious node could prevent other nodes from being configured and entering into the MANET.

(iii) *Address Conflict Threat*. A malicious node can assign a duplicate address to a requester from a possible set of addresses already in use. Thus, it will create, in the DAD process, a blackhole attack of address reply messages (AREP) and lead to an address conflict in the MANET.

(iv) *False Address Conflict Threat*. A malicious node might answer in an unscrupulous way, during the DAD process, using messages AREQ (address request) with false addresses in messages AREP (address reply) that cause conflict with the requester node. Since the victim nodes cannot verify the authenticity of the proposed address, it would have to give up their address and find a new one. The malicious node may change its IP address to execute its attack.

(v) *Denial of Service Threat*. A malicious node could, in an autoconfiguration process, act as a requester and send AREQ messages to multiple initiator nodes simultaneously. Similarly, a malicious node may send many fake DAD messages, causing an overload of traffic.

(vi) *Sybil (Multiples Identities) Threat*. A node illegally claims multiple identities (Sybil node). This node can build a new identity or steal an existing legal node. In general, a Sybil node could demand or assign itself many IP addresses.

(vii) *Negative Reply Threat*. When assigning a new IP address, the approval of all preconfigured nodes is required and an attacker can send a negative response to avoid the entry of the new node.

3. Secure Autoconfiguration

The following are currently the most significant proposals that include secure IP address auto-configuration. The operation of each protocol and what threats they are capable of preventing are explained.

Wang et al. [11] propose a scheme of secure IP address auto-configuration for MANETs, which binds each IP address with a public key allowing each node to authenticate itself into the network and thus prevent spoofing identity and other attacks. The following are considered as the four main security threats surrounding MANET auto-configuration: *address spoofing*, *false address conflict*, *address space exhaustion* and *negative reply threats*.

Identity authentication tries to avoid these threats and this paper proposes to relate every IP address to a public key by means of a one-way hash operation;therefore the owner node of a IP address must use the correspondent key public in order to be authenticated by the network of a unilateral way.

It initiates from the following assumption: the MANET is a network with completely private IP addresses. Therefore, all 32 bits (IPv4) or 131 bits (IPv6) can be used to address nodes in the MANET.

In general, in the proposed scheme, node A, which wants to participate in an existing MANET or start a new one, must first randomly generate a key pair (one public and one private) and one secret key. In the second instance, node A calculates a hash of 32 bits for IPv4 or 131 bits for IPv6.

After calculating the hash value, the node in question temporarily uses the resulting value as its IP address, starts a timer, and broadcasts a *duplicate address probe* (*DAP*) message [13] used to check duplicated addresses on the network.

If a node (node B) configured within the MANET, where Node A wants to enter, finds that the IP address contained in the DAP message issued by node A is equal to it, then it must verify the authenticity of the DAP message. First, node B must check that the IP is equal to the resulting hash of the received public key. Secondly, node B verifies the signature of node A, if it finds that such a signature is correct, then node B checks if public key of node A is equal to it and finally verifies the decryption function. If at least one of the last two checks is not fulfilled, it can be confirmed that there has been an address spoofing attack and therefore node B sends an *address conflict notice* (*ACN*) message via broadcast and discards the received DAP message.

Node A, in turn, waits as long as configured in an internal timer. If it does not receive an ACN message, it assumes that the IP is not in use and permanently assigns the address. If instead it receives an ACN message from some node, before starting the process again, it must verify the authenticity of the ACN message received and the signature of the node issuing the ACN. If these checks are correct, node A is safe that the IP address is assigned to another node and must start the procedure to generate a new pair of public/private keys and secret key; otherwise node A simply discards the ACN message and thus prevents *false address conflict* and *negative reply attacks*.

It is clear that the proposed methodology in the auto-configuration process forces a potential attacker to find, before launching an attack, the public key for which the hash function result is equal to the IP address of the victim, since the controls in the nodes include verification of the identity of the sender node. This process must be applied for each message sent; however the protocol clearly controls address conflict, negative reply, and address spoofing attacks but does not counteract the address exhaustion attack since it does not have a way to specify which node is given which IP addresses, allowing one node to repeat the process as often as desired. This process should be subject to an ACN message which certifies that the node will repeat the process because of IP address duplication.

Buiati et al. [12] propose a secure model for auto-configuration in MANET, based on a distributed and self-organizing certificate system, and also include intrusion detection techniques to improve its safety. The proposed model is built on the protocol DCDP [14] with the improvements proposed by Mohsin and Prakash [15], adopting a collaborative trust model described as "K-out-of-N." So when a new node wants to enter the network, it must earn the trust of K of the N total nodes in the network in order to be accepted into it. To this end, nodes are able to generate certificates with varying degrees of trust. Thus, a distrusted node in the network cannot attack by requesting multiple IP addresses to exhaust them or respond to configuration requests in a malicious manner, as well as allow the implementation of intrusion detection techniques [16].

For the security model, an adversary is defined as any node that produces messages with incorrect auto-configuration protocol information. It then specifies that an adversary can attack the network in two ways: request attacks, where the adversary creates a great number of anomalous messages requesting auto-configuration services, or server attack, where the attacker responds maliciously to requests made by other nodes in the network. In order to avoid these types of attacks, the authors differentiate between trusted and distrusted nodes, avoiding the participation of the latter in the auto-configuration protocol.

Even though there is the possibility that a trusted node is compromised the ability to detect reliable nodes that begin to behave abnormally must be implemented as well. This means that the auto-configuration protocol messages must be authenticated so that an adversary cannot create messages on behalf of another node in the network, being capable of detecting and accusing the adversary nodes. In addition, this detection and accusation system should be implemented collaboratively to prevent an adversary of accusing correct nodes of the network, using the same model "K-out-of-N" explained above.

Authentication of auto-configuration protocol messages is performed using digital signatures, which are built based on digital certificates generated by a distributed certifying authority. This is where the model "K-out-of-N" is applied directly, since, even though every one of the nodes can perform the functions of certifying entity, the entity's private key is split between any subset of K nodes in the MANET. When a new node (one that has not been previously connected to the network) wants to get a digital certificate to identify itself to the MANET, it must take a temporary IP address to request a digital certificate to his 1-hop neighbours. When the MANET nodes receive this request, they can issue a partly signed certificate, depending on the policies established, and send it to the requesting node. After receiving K different certificates, the new node has the ability to build a full certificate and begin the auto-configuration process, discarding the temporary IP. The use of a temporary IP can cause collision problems if the IP is already in use in the network, but it is proposed to use a range of dedicated IPs for this purpose.

The biggest problem in the proposed model is the value of K. A high K value increases security, but reduces the availability of the system because members are less likely to find enough nodes to retrieve the necessary key to the CA. Conversely, if K is small, the availability of the auto-configuration service increases, but the system becomes more vulnerable to attacks by adversaries.

Cavalli and Orset [17] propose a secure auto-configuration protocol adapted to the performance of ad hoc networks, which includes the authentication of the nodes within the network that they will be participating in.

In general, it is intended to satisfy the following items with their secure auto-configuration protocol.

(i) At any time a node must be able to enter or leave the network quickly. Likewise, the network must be able to securely and quickly deliver an IP address to a new node. On the other hand, the abrupt departure of a node must not cause chaos within the network.

(ii) To avoid duplicate IP address conflicts, the protocol must ensure that under no circumstances a node enters the network with its own IP address, but instead the network must be able to deliver the right address to join the MANET.

(iii) The protocol should allow each node to check the veracity of the members of the network to which they belong.

(iv) The protocol should be extremely careful with denial of service. For example, it must not allow a malicious node from monopolizing all IP addresses on the network.

The protocol, in addition to satisfying the described requirements above, wants to meet two broad objectives: the first is to provide a mechanism for IP address auto-configuration for nodes belonging to an ad hoc mobile network, optimizing resources such as bandwidth and time, and the second objective is to allow public key exchange between nodes within the network to ensure the authentication.

The proposed protocol ensures safe IP address auto-configuration including the management of public keys for authentication, which allows avoiding the spoofing attack. However one of its greatest failings is that it neither provides nor supports merging networks or prevents malicious behaviour of network participants after these have been authenticated; among these the denial of service attack is worth mentioning since, for example, a malicious node can authenticate n successive times with n different identities in order to exhaust the available addresses; and another form of attack is that the malicious node refuses to authenticate incoming nodes.

According to Hu and Mitchell [18] the problem with auto-configuration protocols is that their behaviour depends on the correct behavior of all nodes involved. Three attacks are then identified. In the first, a malicious node acts as initiator, assigning duplicate addresses to the requester and sending address assignment messages for nodes that do not exist, effectively reducing the number of addresses available for new valid nodes. The second attack consists of a node acting as a requester, by sending requests for address assignment to multiple initiators, collapsing the network due to broadcast messages generated by the latter in search of a valid IP address. For the third attack, a malicious node can respond to all messages generated by an initiator that tries to find an available IP address, denying access of new nodes to the network.

The proposed solution involves the selection of a method to calculate a "trust value" that is just the level of trust from one node to another, which decreases or increases depending on whether the behaviour of a node is malicious or not, respectively. Then, each node must maintain a list of the levels of trust it has for other nodes. It is possible that different nodes can have different trust limits, depending on security policies. In addition, each node must maintain a blacklist, to which it adds the nodes that do not meet the trust limit, in order to ignore all messages sent by them, except to enable it to recalculate the trust values for these nodes.

When a new node joins the network, it broadcasts a message looking for neighbours, including its trust limit. The nodes receiving this message will respond with a message containing a list of nodes that meet this level of trust, so the new node is able to choose a reliable initiator node. For this model to be fulfilled, the number of malicious nodes needs to be less than the number of normal or valid nodes.

In this way, each time a node receives any information from another node in the network, either as part of the initialization of a new node, collision detection or another process of the auto-configuration protocol, the node first calculates the trust value for the node that sent the message. If this value is below the threshold, the node is added to the black list and a message of suspected malicious node is sent. The nodes receiving this message will act in the same way as the first node, and if they find that the node that sent the message of suspected malicious node has a sufficient trust level, the trust value will be calculated for the suspected node, thus ensuring that only reliable nodes are part of the network. Hu and Mitchell [18] propose a process for calculating the trust level and mention other methods [19, 20].

In the analysis of the trust model, only nodes that consistently behave maliciously are noted. That is, those malicious nodes whose only interest is to affect the calculation of trust values of other nodes are not taken into account and they remain as a weakness in the proposal. Other weaknesses in the proposal are caused by the lack of guarantees against Sybil attacks, where a node uses multiple identities in a fraudulent manner, and against identity theft attacks.

Taghiloo et al. [21] propose the Virtual Address Space Mapping protocol (VASM), where nodes are classified into four categories.

(i) *Allocator*: maintain the address space. They assign new addresses to nodes that join the network.

(ii) *Initiator*: intermediate nodes between the Allocator and the Requester node that exchange all messages between them.

(iii) *Requester*: a new node that needs to get an IP address in order to join the network.

(iv) *Normal*: all the other nodes.

According to this protocol, when a new node joins the ad hoc network, it sends a single hop message called *INITIATOR_SEARCH* to find an Initiator node. If there is no response for this message, the node assumes that it is the only node in the network and begins the network creation process. If the new node gets more than one answer, it selects the sender of the first packet that arrives as an Initiator and sends it an address request. The main task of the Initiator is to get a new IP address from its Allocator and assign it to the requesting node.

In this protocol, each network has at least one Allocator. Each Allocator contains an address space used to assign unique addresses to new nodes as added. The method by which nodes are chosen as an Allocator and how the address space is assigned are the main tasks of the protocol. Similarly, to generate a unique IP address, one Allocator can create another Allocator on the network to balance traffic loads. Each Allocator has a list of all Allocators defined in the network.

The security mechanism for auto-configuration [22] is based on an approach of zero knowledge. This approach only requires a one-way hash function and a seed value, which can be generated randomly. The proposal first establishes a connection between two nodes A and B to exchange information using a cryptographic function on a one-way hash function applied on the seed and in conjunction with a large random number and a secret cryptographic key known by both nodes. Each of the protagonists of the communication carries out the cryptographic operation only the first time and sends the result of the operation and the random number to the other, thus avoiding a man-in-the-middle attack.

For subsequent authentications of both nodes, the value of the seed is increased by one at a time, and the hash function value is calculated on the original value of the seed and the new value increased. The value returned by applying the hash function is sent to the node pair that is being communicated. A node applies the hash function. If the value obtained is equal to the value received the first time, the node is authenticated correctly. For the next communications, the seed value must be incremented by one and the previously explained steps are repeated.

Zhou et al. [23] propose a solution in order to manage the public key of an incoming node, which must be distributed while the secure auto-configuration takes place. Otherwise, a malicious node can impersonate the new node that is registered or that distributes the public key. The SA-PKD achieves the goals of the uniqueness of address allocation and the secure distribution of public key.

It is assumed that the work environment is a densely connected MANET with multiple paths between nodes. If there are malicious nodes in the path between the new node and each of the members, the proposed scheme uses multi-hop broadcast to distribute the information encrypted and signed. Each node checks the forwarded packets to detect the modification of messages.

When a malicious node is placed between a new node and a member of the MANET, it is assumed that there is another good node as a neighbour, and if the malicious node modifies the control message, this node can move or increase the transmission power, sending the message again to try to reach the nodes that lie beyond the malicious one.

If the malicious node deletes the control message, the good node will interpret that the malicious node has left the network or moved away. If there is more than one path between the new node and the MANET member, the message can reach its destination through a different path. If there is a single path, the MANET member will not receive the message because the malicious one interposes and deletes it. The proposal uses the HELLO messages in the routing protocols to help the good node identify the malicious behaviour of the attacker, allowing it to move or increase the transmission power to forward the control message.

4. Conclusions

The insertion of new nodes in a MANET during the auto-configuration process can generate new threats due to the instabilities in the behaviour of these kinds of networks, which would create a lack of trust in the transmission of information through them. The current auto-configuration protocols, with the presented vulnerabilities, have not resolved, in their majority, the security problems found during the insertion of new nodes, creating a necessity for proposals that include this last component. However, the research associated to security during auto-configuration of ad hoc networks is a developing field and still needs much work. In this work, a few existing proposals in the field of secure auto-configuration in MANETs are presented, and they were examined against seven of the most common threats that can be found on these kind of networks to determine how secure or vulnerable they are.

Acknowledgments

This work was supported by the Agencia Española de Cooperación Internacional para el Desarrollo (AECID, Spain) through Acción Integrada MAEC-AECID Mediterráneo A1/037528/11. This work was also supported by the Departamento Administrativo de Ciencia, Tecnología e Innovación (COLCIENCIAS, Colombia) through Programa de Recuperación Contingente which funds Project 121545221101 and the Universidad del Norte through the Dirección de Investigaciones, Desarrollo e Innovación (DIDI).

References

[1] T. Clausen and P. Jacquet, "Optimized Link State Routing Protocol (OLSR)," RFC 3626, Internet Engineering Task Force, 2003.

[2] D. Jonson, Y. Hu, and D. Maltz, "The Dynamic Source Routing Protocol (DSR) for mobile Ad Hoc networks for IPv4," RFC 4728, Internet Engineering Task Force, 2007.

[3] C. E. Perkins, E. M. Belding-Royer, and S. Das, "Ad Hoc On-Demand Distance Vector (AODV) routing," RFC 3561, Internet Engineering Task Force, 2003.

[4] R. Ogier, F. Templin, and M. Lewis, "Topology Dissemination Based on Reverse-Path Forwarding (TBRPF)," RFC 3684, Internet Engineering Task Force, 2004.

[5] Y.-C. Hu, D. B. Johnson, and A. Perrig, "SEAD: secure efficient distance vector routing for mobile wireless ad hoc networks," *Ad Hoc Networks*, vol. 1, no. 1, pp. 175–192, 2003.

[6] S. Gupte and M. Singhal, "Secure routing in mobile wireless ad hoc networks," *Ad Hoc Networks*, vol. 1, no. 1, pp. 151–174, 2003.

[7] H. Deng, W. Li, and D. P. Agrawal, "Routing security in wireless ad hoc networks," *IEEE Communications Magazine*, vol. 40, no. 10, pp. 70–75, 2002.

[8] L. J. García Villalba, J. García Matesanz, D. Rupérez Cañas, and A. L. García Matesanz, "Secure extension to the optimised link state routing protocol," *IET Information Security*, vol. 5, no. 3, pp. 163–169, 2011.

[9] L. J. García Villalba, J. García Matesanz, A. L. S. Orozco, and J. D. Márquez Díaz, "Auto-configuration protocols in mobile ad hoc networks," *Sensors*, vol. 11, no. 4, pp. 3652–3666, 2011.

[10] L. J. García Villalba, J. García Matesanz, A. L. García Matesanz, and J. D. Márquez Díaz, "Distributed Dynamic Host Configuration Protocol (D2HCP)," *Sensors*, vol. 11, no. 4, pp. 4438–4461, 2011.

[11] P. Wang, D. S. Reeves, and P. Ning, "Secure address autoconfiguration for mobile ad hoc networks," in *Proceedings of the 2nd Annual International Conference on Mobile and Ubiquitous Systems-Networking and Services (MobiQuitous '05)*, pp. 519–521, July 2005.

[12] F. Buiati, R. Puttini, R. de Sousa Jr., C. J. Barenco Abbas, and L. J. García Villalba, "Authentication and autoconfiguration for MANET nodes," in *Proceedings of the 2nd International Conference on Embedded and Ubiquitous Computing (EUC '04)*, pp. 41–52, 2004.

[13] A. Abdelmalek, M. Feham, and A. Taleb-Ahmed, "On Recent Security Enhancements to Autoconfiguration Protocols for MANETs Real threats and requirements," *International Journal of Computer Science and Network Security*, vol. 9, no. 4, pp. 401–407, 2009.

[14] A. Misra, S. Das, A. McAuley, and S. K. Das, "Autoconfiguration, registration, and mobility management for pervasive computing," *IEEE Personal Communications*, vol. 8, no. 4, pp. 24–31, 2001.

[15] M. Mohsin and R. Prakash, "IP address assignment in a mobile ad hoc network," in *Proceedings of the Global Information GRID—Enabling Transformation through 21st Century Communications (MILCOM '02)*, pp. 856–861, Anaheim, Calif, USA, October 2002.

[16] R. S. Puttini, J. Marc Percher, L. M. Mé et al., "A Modular Architecture for Distributed IDS in MANET," in *Proceedings of the International Conference on Computational Science and Its Applications: Part III*, pp. 91–113, Montreal, Canada, May 2003.

[17] A. Cavalli and J.-M. Orset, "Secure hosts auto-configuration in mobile Ad hoc networks," in *Proceedings of the 24th International Conference on Distributed Computing Systems Workshops*, pp. 809–814, Hachioji, Tokyo, March 2004.

[18] S. Hu and C. Mitchell, "Improving IP address autoconfiguration security in manets using trust modelling," in *Proceedings of the 1st International Conference on Mobile Ad-Hoc and Sensor Networks*, vol. 3794 of *Lecture Notes in Computer Science (LNCS)*, pp. 83–92, Wuhan, China, December 2005.

[19] C. Huang, H. P. Hu, and Z. Wang, "Modeling time-related trust," in *Proceedings of the Grid and Cooperative Computing Workshops*, vol. 3252 of *Lecture Notes in Computer Science (LNCS)*, pp. 382–389, October 2004.

[20] A. A. Pirzada and C. McDonald, "Establishing trust in pure Ad-Hoc networks," in *Proceedings of the 27th Australasian Conference on Computer Science*, pp. 47–54, Dunedin, New Zealand, 2004.

[21] M. Taghiloo, M. Dehghan, J. Taghiloo, and M. Fazio, "New approach for address auto-configuration in MANET based on virtual address space mapping (VASM)," in *Proceedings of the 3rd International Conference on Information and Communication Technologies: From Theory to Applications (ICTTA '08)*, pp. 1–6, Damascus, Syria, April 2008.

[22] M. Tajamolian, M. Taghiloo, and M. Tajamolian, "Lightweight secure IP address auto-configuration based on VASM," in *Proceedings of the International Conference on Advanced Information Networking and Applications Workshops (WAINA '09)*, pp. 176–180, May 2009.

[23] H. Zhou, M. W. Mutak, and L. M. Ni, "Secure autoconfiguration and public-key distribution for mobile ad-hoc networks," in *Proceedings of the IEEE 6th International Conference on Mobile Adhoc and Sensor Systems (MASS '09)*, pp. 256–263, Macau, China, October 2009.