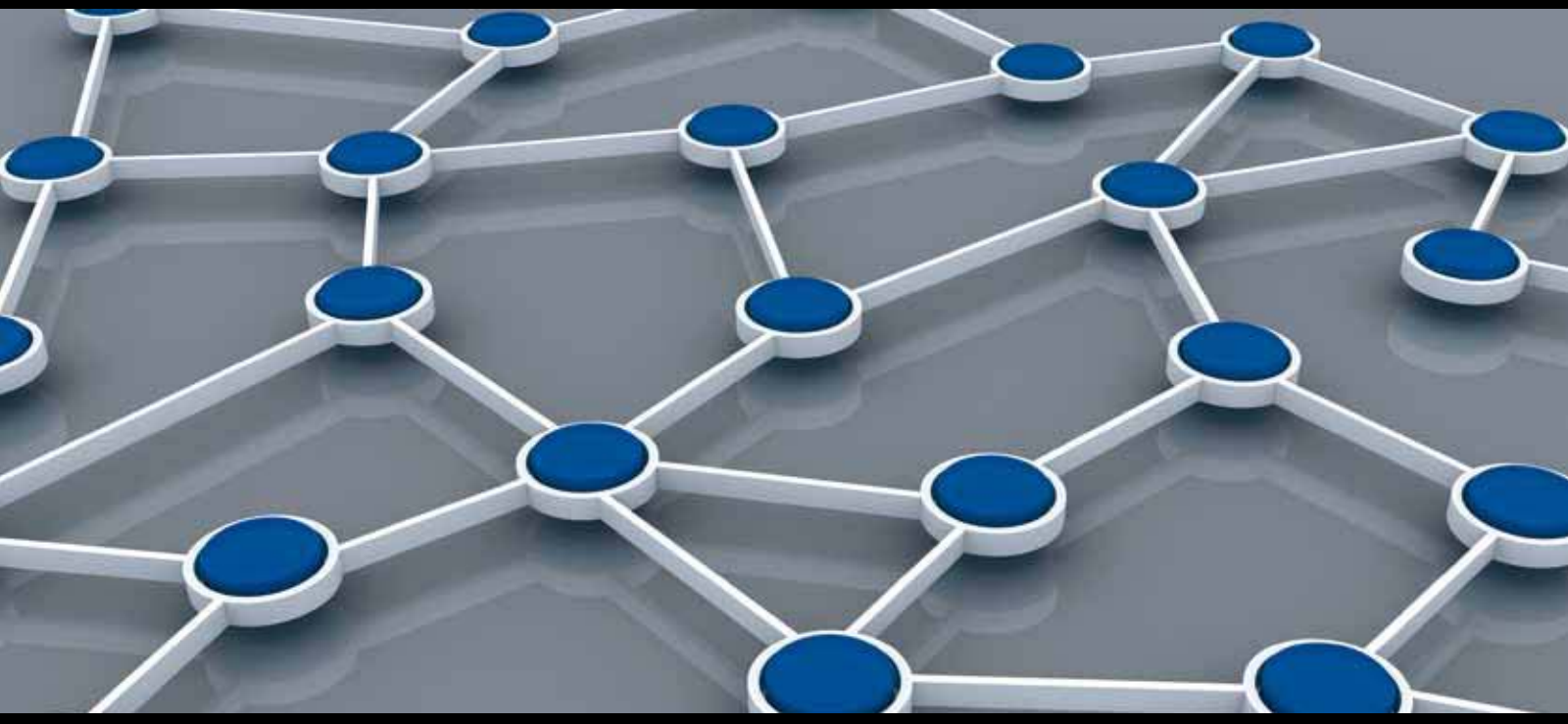
International Journal of Distributed Sensor Networks

Ubiquitous Sensor Networks and Its Application

Guest Editors: Tai-Hoon Kim, Wai-Chi Fang, Carlos Ramos, Sabah Mohammed, Osvaldo Gervasi, and Adrian Stoica



Ubiquitous Sensor Networks and Its Application

International Journal of Distributed Sensor Networks

Ubiquitous Sensor Networks and Its Application

Guest Editors: Tai-hoon Kim, Wai-Chi Fang, Carlos Ramos, Sabah Mohammed, Osvaldo Gervasi, and Adrian Stoica

Editorial Board

Contents

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 310271, 3 pages
doi:10.1155/2012/310271

Editorial
Ubiquitous Sensor Networks and Its Application

Tai-hoon Kim,[1] Wai-Chi Fang,[2] Carlos Ramos,[3] Sabah Mohammed,[4] Osvaldo Gervasi,[5] and Adrian Stoica[6]

[1] *GVSA and University of Tasmania, Australia*
[2] *National Chiao Tung University, Hsinchu, Taiwan*
[3] *GECAD, ISEP, Porto, Portugal*
[4] *Lakehead University, Orillia, ON, Canada*
[5] *University of Perugia, Perugia, Italy*
[6] *NASA Jet Propulsion Laboratory, Los Angeles, CA, USA*

Correspondence should be addressed to Tai-hoon Kim, taihoonn@paran.com

Received 18 April 2012; Accepted 18 April 2012

This special issue contains 24 articles among totally 125 papers accepted in the UCMA 2011 and AST 2011 after submission. Achieving such high-quality papers would have been impossible without the huge work that was undertaken by the International Program Committee members and external reviewers. We take this opportunity to thank them for their great support and cooperation.

Ubiquitous sensor networks and its application are emerging rapidly as an exciting new paradigm to provide reliable and comfortable life services. The ever-growing ubiquitous sensor networks and its application will provide an intelligent and ubiquitous communication and network technology for tomorrow. That is, the UCMA have emerged rapidly as an exciting new paradigm that includes ubiquitous, grid, and peer-to-peer computing to provide computing and communication services at anytime and anywhere. In order to realize the advantages of such services, it is important that intelligent systems be suitable for UCMA.

In "*Logic macroprogramming for wireless sensor networks*," authors evaluated Sense2P analytically and experimentally. Their evaluation result indicates that Sense2P successfully realizes the logic macroprogramming concept while consuming minimal energy as well as maintaining completeness and soundness of the answers.

Authors propose a traffic rerouting scheme in wireless communication infrastructure for islanded microgrid, in "*Traffic rerouting strategy against jamming attacks in WSNs for microgrid*." Authors determine disjoint multiple paths as candidates of a detour path and then select the detour path among the candidates in order to reduce the effect of jamming attack and distribute traffic flows on different detour paths.

The paper "CA_{5W1H} *Onto: ontological context-aware model based on 5W1H*" proposes an ontology-based context-aware modeling technique, along with a relevant framework, in order to enable efficient specification of contextual information and, thereby, further to provide intelligent context-aware services for context management and reasoning.

In "*An energy efficient localization-free routing protocol for underwater wireless sensor networks,*" authors propose an energy efficient routing protocol, named EEDBR (energy-efficient depth-based routing protocol) for UWSNs. EEDBR utilizes the depth of sensor nodes for forwarding data packets. Furthermore, the residual energy of sensor nodes is also taken into account in order to improve the network life time.

In "*A novel coverage enhancement algorithm for image sensor networks*," based on virtual potential field, the Paired tangent point repulsion for nonboundary sensor nodes

and fuzzy image recognition for boundary sensor nodes realize the enhancement of perspective coverage, together with LRBA, MBAA, and mixed superposition algorithm for rotation angle adjustment.

In "*A communication framework in multiagent system for islanded microgrid*," authors design a communication framework to control and operate distributed sources and loads in the islanded microgrids. The framework reliably delivers microgrid control frame between agents by employing wireless mesh network as an advanced topology of the wireless sensor network.

The paper "*Real-time train wheel condition monitoring by fiber Bragg grating sensors*" describes a real-time system to monitor wheel defects based on fiber Bragg grating sensors. Track strain response upon wheel-rail interaction is measured and processed to generate a condition index which directly reflects the wheel condition.

The paper "*Energy-efficient fire monitoring over cluster-based wireless sensor networks*" proposes an energy-efficient fire monitoring protocol over cluster-based sensor networks. This proposed protocol dynamically creates and recognizes the sensor network cluster hierarchy according to the direction of fire propagation over the sensor network clusters.

The purpose of the paper, "*Networked Electronic Equipments using the IEEE 1451 Standard: VisioWay, a Case Study in the ITS Area*" is to analyze the integration of electronic equipments into intelligent road-traffic management systems by using the smart transducer concept. An automated video processing sensor for road-traffic monitoring applications is integrated into an ITS network as a case study.

In "*The vegetable freshness monitoring system using RFID with oxygen and carbon dioxide sensor*," authors use a sensor for monitoring gases and combine the sensor with an RFID tag. The RFID system is relatively easy to manage. With this combined system, we estimated the freshness of vegetables.

The paper "*An adaptive system supporting collaborative learning based on a location-based social network and semantic user modeling*" presents an adaptive e-learning system which supports collaborative learning based on a location-based social network and semantic modeling. In the system, a social network among e-learning learners is dynamically constructed on the basis of the location information of learners using GPS sensors for collaborative learning.

The paper "*Design of an effective WSN-based interactive u-learning model*" presents a model of an effective and interactive ubiquitous learning environment system based on the concepts of ubiquitous computing technology that enables learning to take place anywhere at any time.

In "*A local world evolving model for energy-constrained wireless sensor networks*," authors aim at improving the interactions among sensor nodes and present a heterogeneous local-world model to form large-scale wireless sensor networks based on complex network theory.

In "*An energy supply system for wireless sensor network nodes*," the overall system structure, the function modules design, and the performance testing analysis are illuminated in detail. Experimental results reveal that this energy supply system can significantly improve power within the wide bands by the active piezoelectric energy harvesting technology and enable wireless sensor network nodes to operate normally.

In the paper "*Reliable latency-aware routing for clustered WSNs*," a unique latency sensitive reliable routing protocol for WSNs has been proposed. This protocol uses the concept of hotlines (high-reliable links) and also utilizes alternative routes to reduce the number of hops from the source to the sink.

The paper "*An optimization scheme for M2M-based patient monitoring in ubiquitous healthcare domain*" performs optimization scheme movement coordination technique and data routing within the monitored area. A movement tracking algorithm is proposed for better patient tracking techniques and aids in optimal deployment of wireless sensor networks.

In the paper "*A dynamic traffic-aware duty cycle adjustment MAC protocol for energy conserving in wireless sensor networks*," a dynamic traffic-aware MAC protocol for energy conserving in wireless sensor networks is proposed. The proposed MAC protocol can provide better data transmission rate when sensors with high traffic loading.

In "*Grid-based predictive geographical routing for inter-vehicle communication in urban areas*," authors propose a grid-based predictive geographical routing (GPGR) protocol, to which overcomes these problems. GPGR uses map data to generate a road gird and to predict the exact moving position of vehicles in during the relay node selection process.

In "*A wireless sensor network for precise volatile organic compound monitoring*," A variety of methods have been developed to monitor VOC concentration in hazardous sites. The methods range from calculation to measurement, and from point measuring to remote sensing.

The paper, "*The construction of inference engine for meaningful context and prediction based on USN environment*" proposes to design for this through application of context inference of USN (ubiquitous sensor network) and inference production rules for context inference engine of wetland management system by using JESS.

In the paper "*An ACOA-AFSA fusion routing algorithm for underwater wireless sensor network*," a novel ACOA-AFSA fusion algorithm for UWSN routing protocol has been presented. It is a useful routing algorithm for underwater sensor networks owing to its local acknowledge and global view offered by ACOA and AFSA, respectively.

In "*Improved virtual potential field algorithm based on probability model in three-dimensional directional sensor networks*," authors propose a 3D directional sensor coverage-control model with tunable orientations. Besides, a novel criterion for judgment is proposed in view of the irrationality that traditional virtual potential field algorithms brought about on the criterion for the generation of virtual force.

In the paper "*A network coding based rarest-first packet recovery algorithm for transmitting geocast packets over hybrid*

sensor-vehicular networks" a network coding based rarest-first packet recovery algorithm for transmitting geocast packets over hybrid sensor-vehicular-networks is proposed.

The paper, "*Anchor-node-based distributed localization with error correction in wireless sensor networks*" proposes a scheme to enhance localization in terms of accuracy and transmission overhead in wireless sensor networks. This scheme starts from a basic Anchor node-based distributed localization (ADL) using grid scan with the information of anchor nodes within two-hop distance.

Tai-hoon Kim
Wai-Chi Fang
Carlos Ramos
Sabah Mohammed
Osvaldo Gervasi
Adrian Stoica

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 472986, 6 pages
doi:10.1155/2012/472986

Research Article

The Vegetable Freshness Monitoring System Using RFID with Oxygen and Carbon Dioxide Sensor

Ki Hwan Eom, Min Chul Kim, SeungJoon Lee, and Chang won Lee

Department of Electronics and Electrical Engineering, Dongguk University, 26 Pil-dong 3-ga, Jung-gu, Seoul 100-715, Republic of Korea

Correspondence should be addressed to Ki Hwan Eom, kihwanum@dongguk.edu

Received 4 November 2011; Revised 11 April 2012; Accepted 11 April 2012

Academic Editor: Tai Hoon Kim

This paper proposes an oxygen and carbon dioxide concentration monitoring system for freshness management based on radio frequency identification (RFID). Freshness can be checked by various factors including humidity, temperature, oxygen, and carbon dioxide. This paper focuses on oxygen and carbon dioxide. The concentrations of these two gases are related to freshness and affect the food. We use a sensor for monitoring these gases and combine the sensor with an RFID tag. The RFID system is relatively easy to manage. With this combined system, we estimated the freshness of vegetables.

1. Introduction

The vegetable has relatively short product availability period. When we buy the vegetable, we want to check the freshness criteria. However, there is no such a system that can check the freshness of vegetables, so people just inspect visually. If the vegetable goes beyond the expiration date, people will throw it away, so it causes huge waste of money and may threat customers' health. There will be needed certain freshness monitoring system for both customers and seller to save money and health.

Oxide and carbon dioxide are needed for organisms to survive. Microorganisms absorb oxygen and emit carbon dioxide as food spoil [1]. The respiration of food in package also affects food freshness. We believe freshness can be estimated by monitoring the levels of oxygen and carbon dioxide. Freshness is affected by many factors including moisture and temperature, oxygen. Until now, the research of freshness was limited by temperature and humidity, and temperature and humidity have been managed by sellers themselves. Therefore there should be more research on oxide and carbon dioxide for checking freshness factors.

This paper proposes oxygen and carbon dioxide concentration monitoring system or freshness management based on RFID. The proposed system uses two sensors to measure oxygen and carbon dioxide for monitoring these two gases. The oxygen sensor's type is galvanic cell. This sensor does not need power supply device, so we can easily design the circuit for monitoring system. The RFID is very useful for various applications because this system is very small, uses non- or very small capacity battery, and is easy to use its application [2–5]. Thus, the proposed system uses RFID with two sensors, so freshness can be checked more convenently and faster.

In the next chapter, we will discuss the system proposed with circuit and block diagram. And finally, Chapter 3 concludes the paper.

2. Proposed Oxygen and Carbon Dioxide Monitoring System

Figure 1 shows the RQ (respiration quotient) of mature green mume (green plum) in packages with different transmission rates of oxygen and carbon dioxide. Table 1 shows the detailed data for Figure 1. This RQ links the oxygen consumption rate with the carbon dioxide creation rate. This happens as food "breathes." If the RQ is more than 1, food freshness will decrease [6, 7]. This paper proposes an oxygen and carbon dioxide monitoring system to check freshness.

In this paper, we use sensors for monitoring of vegetable freshness. So this paper selects the sensors that operate the low temperature and humidity of wide area because the food keeps storage of the low temperature for maintaining

Table 1: Package of different transmission rate of oxide and carbon dioxide.

Films	Real thickness (μm)	Gas transmission rate (RH of 76% @ 25°C) (mL/m^2·day·atm)		Water vapor transmission rate (RH of 100% @ 38°C) (g/m^2·day·atm)
		O_2	CO_2	
LDPE A	18	2,694	9,776	19.81
LDPE B	27	2,142	6,711	17.68
LDPE C	51	1,568	4,580	12.84

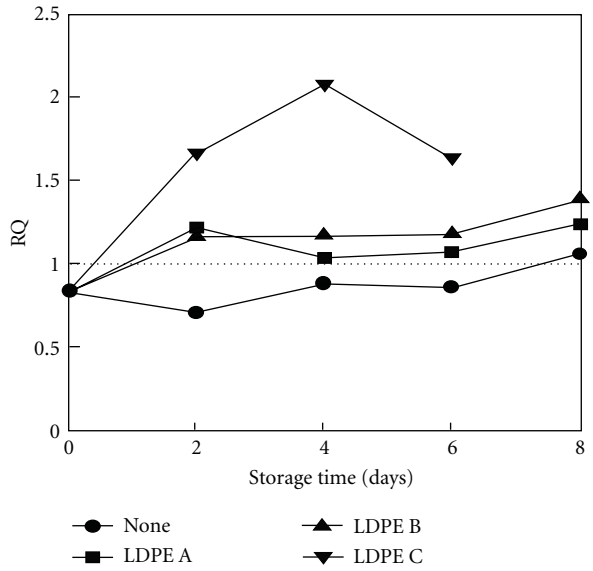


Figure 1: Changes in respiration quotient (RQ).

Table 2: SS1118 specifications.

Content	Specification
Measurement range	0~100% oxygen
Output signal	6 ± 1 mV in RF of 40% 25°C
Temperature range	−10~50°C
Operating humidity	0~99% RH

Table 3: NAP-21A specifications.

Content	Specification
Voltage supply	D.C. 1.8 ± 0.18 V
Measurement range	0~100% carbon dioxide
Output voltage	0~20 mV
Temperature range	−10~50°C
Operating humidity	0~95% RH

freshness. In addition we also check the input and output volt and current of sensors, because these sensors connect the RFID. The RFID's output and input volt and current are very small. Therefore we have to select the sensors that output and input almost appropriate voltage and current at these RFIDs. So this paper selects the oxygen sensor and carbon dioxide sensor at SS1118 and NAP-21A. These sensors are shown in Figure 2. Figure 2(a) selects oxygen sensor (SS1118) and Figure 2(b) selects carbon dioxide sensor (NAP-21A).

The SS1118 oxygen sensor is of galvanic cell type as shown in Figure 3. The galvanic cell type has electrode, and this electrode generates the electric voltage according to oxygen concentration such as Figure 4.

This oxygen gas sensor requires no special preparation or calibration–just plug it into your interface and it is ready to take readings because it just generates the power. So it is very easy to use for connecting with RFID system that requires low power consumption battery. Especially, this sensor offers superior performance over the conventional oxygen sensor in that it is not affected by carbon dioxide, carbon oxide, and nitrogen oxides.

Table 2 is specifications of SS1118. We can see that this sensor is suitable for connecting with RFID and using freshness monitoring system. The range of operating temperature and humidity is suitable at checking the vegetable freshness. Typically the vegetable is stored in low temperature and high humidity. This sensor covers enough the storage environment of vegetable.

We design the circuit for measurement of output voltage and connect with RFID system such as Figure 5. This circuit produces more stable output. We do not need input circuit this sensor. The sensor type is galvanic cell. This type does not require the input power.

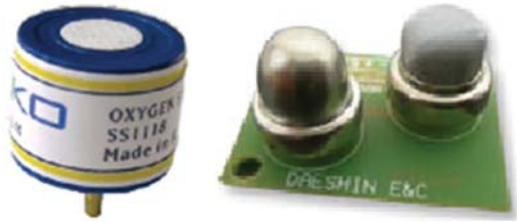
Next, we will check the NAP-21A carbon dioxide sensor. This sensor is thermal conductive type, and it is able to detect a wide range of carbon dioxide gases up to 100%. The thermal conductive type sensor measures heat conductivity according to carbon dioxide concentration.

This sensor is appropriate for our application system according to Table 3. It operates with the very low voltage and temperature irrespective of humidity.

Figure 6 is the suggested circuit, and this sensor is thermal conductive type whose accuracy is somewhat low. We use compensator in this circuit and make the bridge circuit like Figure 6. We can measure and connect the RFID using this circuit.

Finally, the RFID system stores the data in the RFID tag and receives data to RFID reader. This data is result of measurement from the sensors. Thus the system connects between sensors and RFID tag to store the sensing data. So this paper uses the MLX90129 chip for RFID tag because this chip combines a precise acquisition chain for external resistive sensors, with a wide range of interface possibilities.

We connect the oxygen sensor and carbon dioxide sensor like Figure 6. The input and output signal of sensors is



(a) (b)

Figure 2: Using sensors: (a) oxygen sensor (SS1118) and (b) carbon dioxide sensor (NAP-21A).

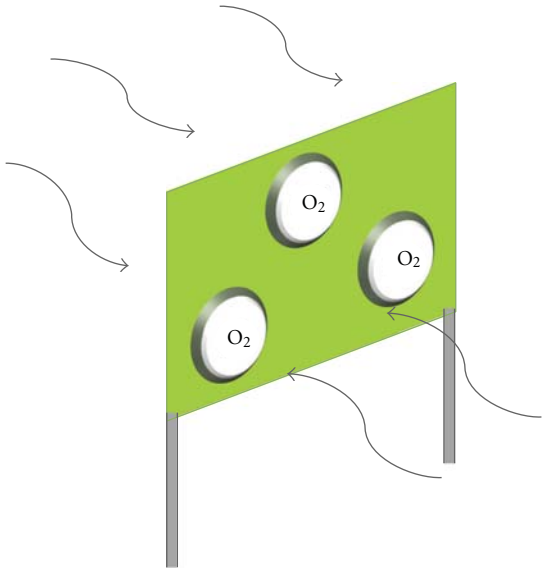


Figure 3: Operation of galvanic cell type sensor.

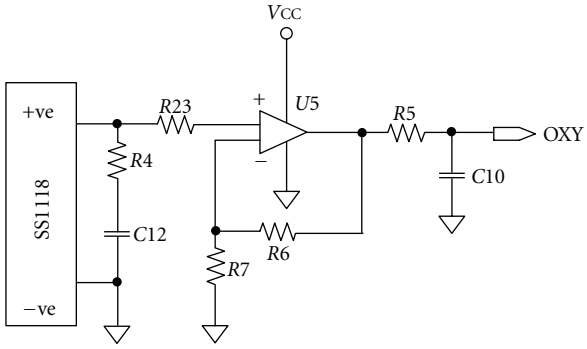
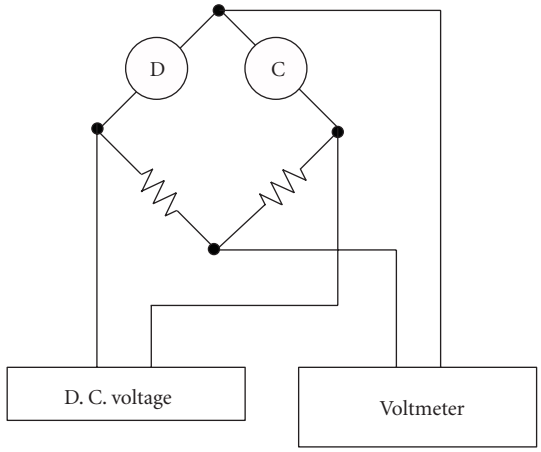


Figure 4: Output circuit design of SS1118.



D: Detector
C: Compensator

Figure 5: Circuit design of NAP-21A.

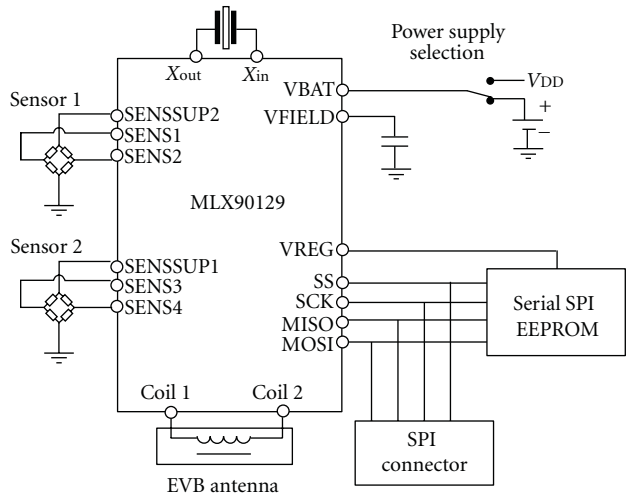


Figure 6: RFID block diagram in proposed system.

included in MLX90129 input and output signal range. So we do not consider adjusting the in-out signal size. This RFID tag senses oxygen and carbon dioxide and saves the collected data to EEPROM. This data will be read to RFID reader when the reader requests the data to the tag. If we use the RFID system, we can find the data easier.

3. Experiments

First we must check the operation circumstance of sensor. The sensor output is a voltage data but we need the data of oxygen and carbon dioxide concentration, so we set the experiment as Figure 7.

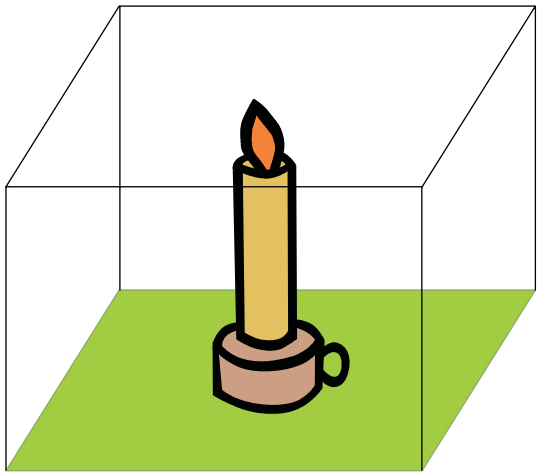


Figure 7: Experiment settings for transformation of the data form voltage to concentration.

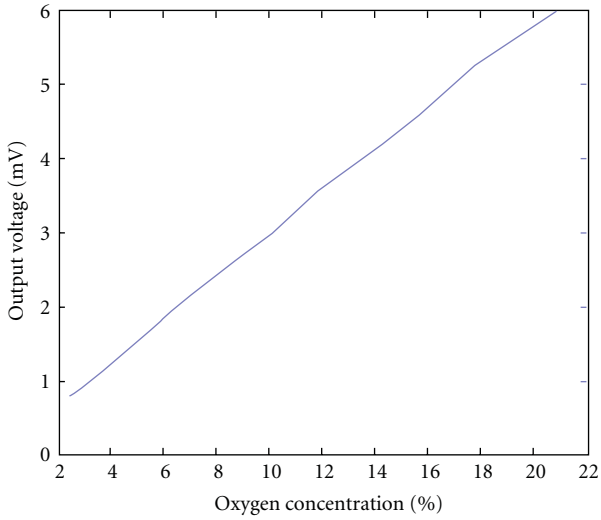


Figure 8: Oxygen concentration versus sensor output voltage.

We use a candle for transformation concentration of oxygen and carbon dioxide (in order to keep the light of candle, candle light consumes the oxygen in the experiment area and increases proportion of the carbon dioxide). We check the sensor data by DMM and compare it with data of oxygen and carbon dioxide concentration instrument.

Figure 8 is result of oxygen concentration versus sensing data. And Figure 9 is result of carbon dioxide concentration versus sensing data. Following this comparative result data, we will estimate the concentration of gases. The result data is almost linear so we can use it easily:

$$\begin{aligned} &\text{oxygen concentration (\%)} = 3.48 * \text{V (mV)},\\ &\text{carbon dioxide concentration (\%)} = 5 * \text{V (mV)}. \end{aligned} \tag{1}$$

Equation (1) is derived by graph of Figures 8 and 9. If we use this equation, we can know the gas concentration to see

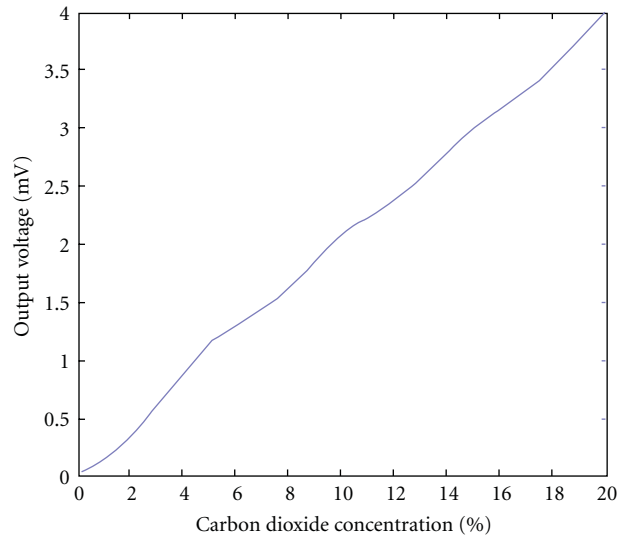


Figure 9: Carbon dioxide concentration versus sensor output voltage

the output voltage. The suggested system outputs the voltage information in RFID reader. So we need to calculate the gas concentration using voltage information.

Finally, we experiment on monitoring of oxygen and carbon dioxide concentration. This experiment uses the tag that is connected with sensors and checks the changed concentration of oxygen and carbon dioxide inside the package of the vegetable which is shown Figure 10. The experimental environment is as follows: temperature: 20°C and humidity: 35% RH. We can find the data easily showing RFID reader data as shown as Figure 11. We use shown data divided by 10000 and the value is mV.

Figure 12 is result of this experiment for a week. We can know that the oxygen and carbon dioxide concentration in package is changed every day through Figure 12. The vegetable in package consumes the oxygen and emits the carbon dioxide. And also the freshness of vegetable is changed. Therefore we can monitor the freshness checking the oxygen and carbon dioxide concentration. Furthermore, using this data, we can easily check and display the freshness with LEDs color as shown in Figure 13.

4. Conclusion

Nearly all organisms need oxygen and carbon dioxide to survive. Food also breathes and gradually spoils. If we observe oxygen and carbon dioxide used to breathe, we can check food freshness. In this paper, we observe these two gases concentrations using sensor. This sensor must have a wide operation range. Vegetables may be stored in low temperature and humidity, so a sensor has to endure this environment and others.

By combining gas sensors and RFID tags, it is relatively easy to monitor vegetable freshness. The proposed system uses RFID tags that get data on oxygen and carbon dioxide concentration. By checking RFID reader, we can track how

Figure 10: The vegetable inside the package for the experiment.

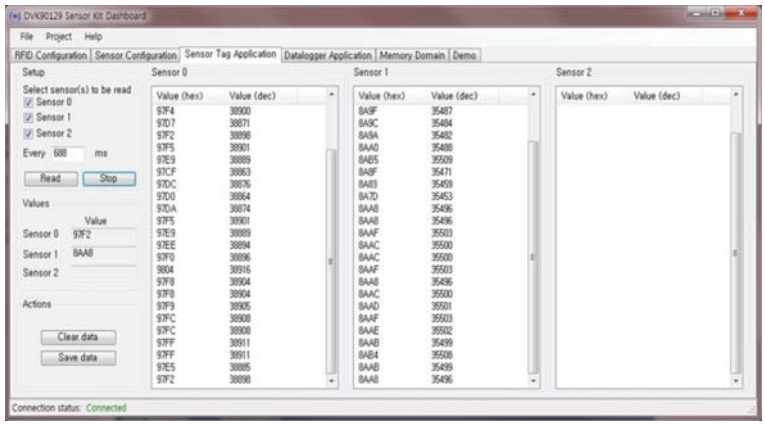


Figure 11: The reading data of RFID reader.

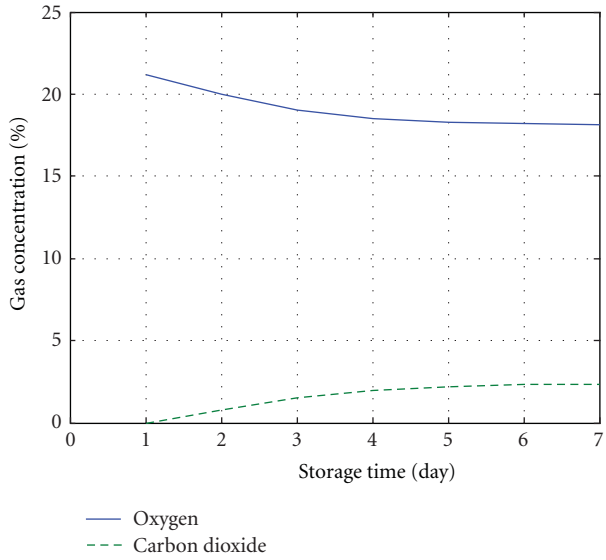


Figure 12: Results of monitoring gases concentration.

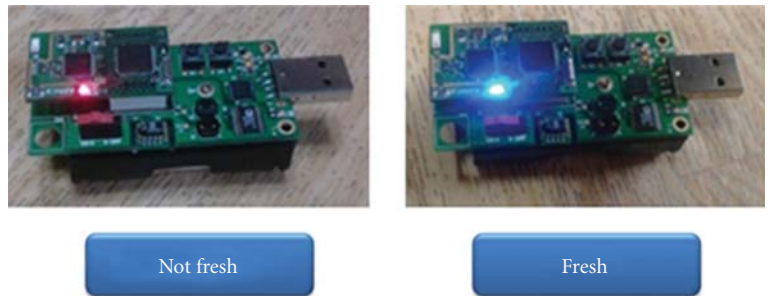


Figure 13: Display freshness using LEDs.

oxygen and carbon dioxide concentrations and vegetable freshness change over time. Furthermore, using this data, we can easily check and display the freshness with LEDs color.

Although this paper offered an initial contribution to combining two gas sensors and RFID tags, a further research could be continued on developing the smart RFID tag that has more sensors to get more precious data on food freshness.

Acknowledgment

This research was supported by the Agriculture Research Center (ARC, 710003-03-1-SB110) program of the Ministry for Food, Agriculture, Forestry and Fisheries, Korea.

References

[1] H. S. Chae, C. N. Ahn, Y. MoYoo et al., "Effect of water uptake rate of chicken on lipid oxidation, color of meat, and microbes of chicken during storage," *Journal of Korea Poultry*, vol. 35, no. 3, p. 247, 2008.

[2] Wikipedia, http://www.en.wikipedia.org/.

[3] K. Finkenzeller, *RFID Handbook*, Wiley, 2010.

[4] H. J. Kim, E. J. Woon, and J. J. Woo, "Cryptanalysis and improvement of an RFID authentication protocol based on private codes," *Korean Institute of Information Technology*, vol. 9, no. 5, pp. 103–110, 2011.

[5] T. W. Ahn, S. C. Jeong, Y. S. Seo, and D. Y. Lee, "Validation evaluation of RFID application for logistics industry—in case of a bonded logistics system," *Korean Institute of Information Technology*, vol. 9, no. 4, pp. 155–166, 2011.

[6] S. H. Cho, J. W. Heo, Y. J. Choi, J. H. Kang, and S. H. Cho, "Effects of grapefruit seed extract and an ion solution on keeping quality of mungbean sprouts," *Korean Journal of Food Preservation*, vol. 12, no. 6, pp. 534–539, 2005.

[7] H. S. Cha, S. I. Hong, J. S. Park et al., "Respiratory characteristics and quality attributes of mature-green mume (Prunus-mumeSieb. et Zucc) fruits as influenced by MAP conditions," *Journal of Korean Society Food Science and Nutrition*, vol. 28, no. 6, pp. 1304–1309, 1999.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 370935, 11 pages
doi:10.1155/2012/370935

Research Article

A Novel Coverage Enhancement Algorithm for Image Sensor Networks

Haiping Huang,[1,2,3,4] Lijuan Sun,[1,2,3] Ruchuan Wang,[1,2,3] and Jing Li[1,2]

[1] *College of Computer, Nanjing University of Posts and Telecommunications, Nanjing 210003, China*
[2] *Jiangsu High Technology Research Key Laboratory for Wireless Sensor Networks, Nanjing University of Posts and Telecommunications, Nanjing 210003, China*
[3] *Key Lab of Broadband Wireless Communication and Sensor Network Technology of Ministry of Education, Nanjing University of Posts and Telecommunications, Nanjing 210003, China*
[4] *College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China*

Correspondence should be addressed to Haiping Huang, hhp@njupt.edu.cn

Received 3 November 2011; Revised 13 March 2012; Accepted 13 March 2012

Academic Editor: Sabah Mohammed

The needs of diverse environmental information introduce the multimedia data into wireless sensor networks. The characteristics of most multimedia information, such as great amounts of data and high-quality requirement for network service, positively affect traditional wireless sensor networks, which also derive various new research areas. This paper focuses on the multimedia image sensor networks and proposes FVPTR (fuzzy image recognition and virtual-potential-field-based paired tangent point repulsion) method to enhance the perspective coverage of network. This approach utilizes fuzzy image recognition method to process the boundary nodes. Aimed at nonboundary nodes, based on potential field theory, it adopts paired tangent point repulsion mechanism, which attempts to obtain the optimal network sensing coverage through the multiple paired achievements between one current node and several target nodes. Combined with FVPTR, some algorithms such as LRBA, MBAA, and mixed superposition algorithm are put forward to single or multiple-time adjustment by the rotation of the direction angle. The results of simulation and all kinds of comparisons show that three-times pairing method enhances the coverage of networks well.

1. Introduction

Wireless senor networks (WSNs) enjoy great applications in traditional fields such as industry, agriculture, military, and environmental monitor. Besides, WSNs also show the superiority in areas of household paradigm, health care, transportation, and so forth. With the occurrence of new applications, the needs for diverse environmental information from users are increasing, and the multimedia information is introduced into wireless sensor networks. The characteristics of most multimedia information, such as amounts of data and high-quality requirement for network service, greatly affect the traditional techniques of WSNs and meanwhile derive some new research points. Deployment and coverage are two typical issues, which not only reflect the perceptive ability of networks to physical world, but also are directly related to the quality of network services [1].

Numerous researches have focused on the coverage issues [2–7]. Jing and Alhussein [5] gave out a coverage model of target points for direction sensor, proposing LPI algorithm to compare with CGA and DGA algorithms. Judging from the simulation results, it could not show significance in coverage enhancement, and it was a preliminary study on directional sensors. Tao et al. [8] changed the coverage-enhancing problem into virtual-potential-field-based centroid points' uniform distribution problem which ignored the effect of border nodes. Mohamed and Hossien [6] proposed PCP protocol based on omnidirectional image sensors and deduced several models, in which the exponential model is an inspiration for our method. Zou and Krishnendu [9] used VFA algorithm to generate the mobile paths for sensor nodes based on virtual potential field and artificially changed the location of each node in accordance with the calculated trajectory.

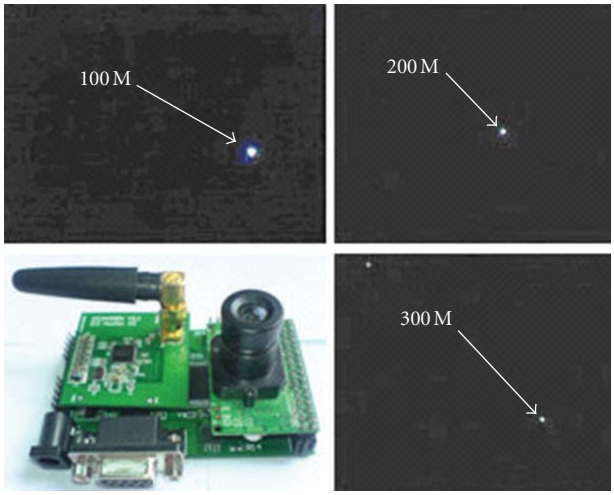


Figure 1: Photos of point light source at 100 m, 200 m, and 300 m away from the image sensor node.

Nonetheless, it is almost impossible to achieve in the case of large-scale deployment.

Image sensor is a typical case of directional sensors. This paper discusses the scenario that stochastically deployed image sensors in limited areas and uses paired tangent point repulsion method based on fuzzy image recognition and virtual potential field theory to improve the coverage of image sensor networks.

2. Preliminary

2.1. Fuzzy Image Recognition. Oztarak et al. [10] demonstrated that image sensors could perceive "video event" once accessing into field of view (FOV) [5, 11]. He utilized joint fuzzy processing method with micro-SEBM (structural and event based multimodal) to compose the mobile trace of "video event" in an image captured by nodes and demonstrated that image sensors had the ability to identify the "video event." Then the specific location of "video event" would be positioned in the current image through scanning it and be clearly expressed by MBR (minimum bounding rectangle) information which could be recorded for further operation.

The frequency and the amounts of data of most monitors and household cameras cannot be afforded by wireless multimedia sensor networks (WMSNs). A specific image sensor node developed according to actual demand would be used as the hardware to test the fuzzy image recognition method in the following experiments.

The node as shown in Figure 1 is equipped with ATmega128 processor, OV7620 image processing chip, and CC2420 communication chip. Moreover, it can work at two modes of 24-bit color and 8-bit grayscale, and sample images at three different pixel resolution ratios such as $88 * 72$, $160 * 120$, and $320 * 240$.

The flashlight with AA battery is used as point light source in this experiment. The node working at the mode of 24-bit color and the pixel of $160 * 120$ takes photographs on the point light source, as shown in Figure 1.

The image sensor node scans the data of photo. It has not distinguished the point light source until scanning the Xth line and the Yth column and then records the value of column in an array named D until the light source disappears from the Xth line. When scanning the $(X + K)$th line, the same operation would be done until the node finishes scanning the whole image. After calculating the average value of D, the average column value (ACV) of "video event" in the current image can be figured out. The specific application of ACV will be described in Section 4.1.

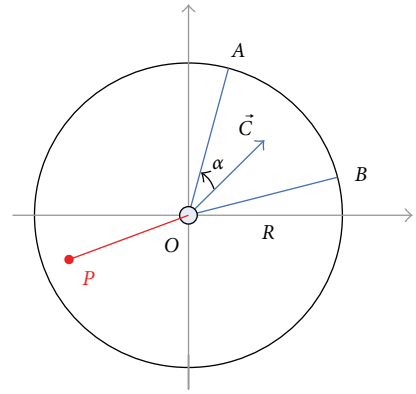


Figure 2: Directional sensing model.

2.2. Directional Sensing Model. In Figure 2, the effective sensing area of directional sensor is fan-shaped region of OAB which can rotate around the point O. The perceptive radius of directional sensor is R, which will be depicted in Section 2.3. $\vec{C}$ is a unit vector that begins at O, points to the center of fan-shaped region, and represents a sensor's direction of effective sensing area named as perceptive direction. By adjusting it, a circular area within R can be completely covered. 2α represents the sensor's FOV, which should be approximately $\pi/3$ for the special image sensor node in this paper according to actual tests.

At a discrete moment, it can be determined whether any point P is covered by a directional node or not by the following:

$$\begin{aligned} \left\|\overrightarrow{OP}\right\| &\le R, \\ \overrightarrow{OP} \cdot \vec{C} &\ge \left\|\overrightarrow{OP}\right\| \cdot \cos\alpha. \end{aligned} \tag{1}$$

If a point P meets the above two conditions simultaneously, it is covered.

There are two concepts that should be distinguished well.

(1)"Effective sensing area": fan-shaped region of OAB shown in Figure 2 can be calculated by measure of area.

(2) "FOV": sensor's field of view is the angle of effective sensing area and can be depicted by the dimension of angle.

2.3. Perceptive Radius Theory. Here, two conceptions of perceptive radius are proposed.

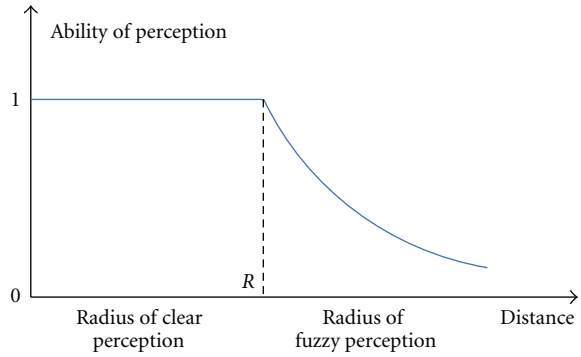


Figure 3: Perceptive ability and perceptive radius.

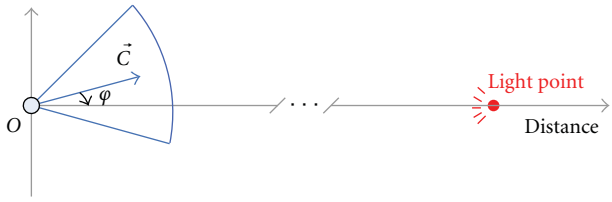


Figure 4: Image sensor node and sensing light source.

(1) "Radius of clear perception": in range of this radius, an image sensor node can clearly identify any object accessing into its FOV. The scope of clear perceptive radius is $[0, R]$. The perceptive ability of an image sensor remains constant in the range of clear perceptive radius as shown in Figure 3.

(2) "Radius of fuzzy perception": in range of this radius, an image sensor node cannot clearly identify the object accessing into its FOV. The requirements of environmental monitoring cannot be well met, but the node can respond to special "video event." Strong light source is such a typical case of special "video event." The scope of fuzzy perceptive radius is $[R, \infty]$. As depicted in Figure 3, node's sensing capacity exponentially declines within the range of fuzzy perceptive radius. Nevertheless, the node's response to special "video event" should not be greatly affected.

One case is presented as follows: there is an image sensor at the starting point of O. Meanwhile, a strong light source is set at an infinite distance away from starting point as shown in Figure 4. The light source lies in the range of the node's fuzzy perceptive radius. The sensor cannot clearly identify the targets near the light source, that is, it also cannot monitor the environment there. But the "video event" of strong light source can be perceived just as the experiment described in Section 2.1.

The angle between $\vec{C}$ and the vector which begins at O and points to light source is named as φ. It is in line with the above-mentioned case when φ is roughly in the range of $[-\pi/6, \pi/6]$. In this situation, an image sensor node can find out the specific location of "video event" by scanning the current captured image. Inspired by model of micro-SEMB [10], we propose the video event search algorithm using ACV, which will be discussed in Section 4.1.

2.4. Virtual Potential Field. Introducing virtual potential field into WSNs originates from its application in obstacle avoidance [12, 13]. Virtual potential field, due to the simplicity and the characteristic of real time, has been introduced into the coverage problem in WMSNs [8]. In virtual potential field, each node can be considered as a virtual charge, suffers the virtual force from other nearby nodes, and gets the trend of moving towards the region of less node density in networks.

Tao et al. [8] figured out that the focus of each sensor node's effective sensing area would rotate around the node while suffering virtual force from neighboring nodes. Resultant force from all neighboring nodes within the scope of effective communication should be taken into account, which inevitably increased the difficulty of force analysis and the complexity of algorithm. So before FVPTR is executed, the process of pairing between two adjacent nodes should be achieved. The proposed method ignores the influence of most nearby nodes, only focusing on adjusting the two paired nodes' perceptive direction angles in terms of the virtual force between them as reducing the complexity of computation as possible. Moreover, the simulation results prove that FVPTR indeed enhances the perspective coverage of networks.

3. Framework

The fuzzy image recognition and virtual potential field methods are specifically introduced in this section.

3.1. Initialization. Image sensor nodes are randomly and uniformly distributed in a monitored region. Some typical coverage enhancing algorithms often assume that nodes can estimate their locations. Equipping every node with GPS will bring some of the cost, so localization is typically performed by estimating distances between neighboring nodes. In this paper, we focus on application in which localization is unnecessary and possibly infeasible. So we consider that nodes are unaware of their locations.

Strong light source, that causes the "video event," is set in the center of each boundary in the region. Unfortunately, the coverage of networks is unsatisfactory after initial deployment, and there exists large amounts of "overlapped regions." "Overlapped region" can be defined as a region that is covered by more than two nodes effective sensing area at the same time.

In fact, how to enhance the coverage and reduce the "overlapped region" has already been the same issue in this paper.

3.2. Virtual Force. The potential field is a model of electrostatic field in physics. Each node can be seen as a point charge in the electrostatic field and has the same energy. In other words, these nodes are identical with the same type of charge and the same equivalent electricity. Taking into account the nature of the homosexual repulsion and heterosexual attraction, it is reasonable to suppose that the virtual force in this potential field is repulsion between two homogeneous

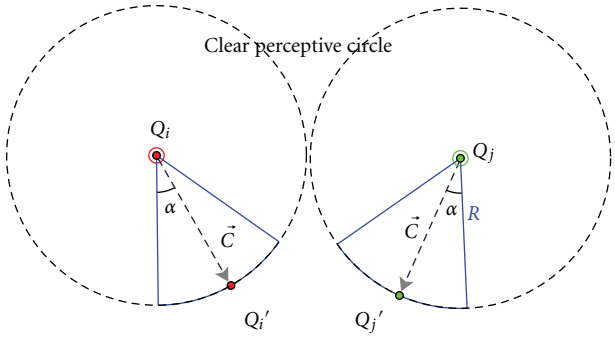


Figure 5: Nearby unpaired nodes.

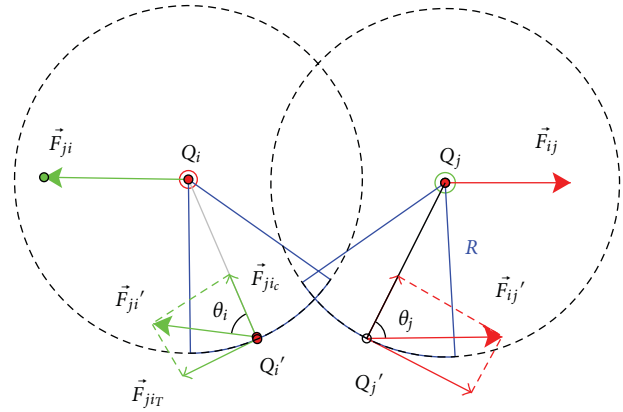


Figure 6: Force analysis of paired nodes.

nodes. But the repulsion is different from general repulsion between charges.

There are two image sensor nodes, whose clear perceptive radius is R, named as Q_i and Q_j, in Figure 5. The two circles with the radius R, whose center is Q_i and Q_j respectively, are called clear perceptive circle.

Not only the data quantity of WMSNs is larger than that of normal WSNs, but also the integrity of multimedia data is more important than that of scalar sensing data. In order to ensure the quality of communication while avoiding frequent data loss or interruption, reducing the actual distance between nodes becomes a valid method.

Effective communication radius of one node marked with R_C is twice that of the radius of node's clear perceptive radius R.

The Euclidean distance between Q_i and Q_j is defined as $D_{ij} = \|\overrightarrow{Q_iQ_j}\|$. When D_{ij} is beyond R_C, the repulsion force between two adjacent nodes would be so tiny that it almost tends to be zero. In this situation, two nearby nodes can independently exist in the respective virtual potential field and ignore the effect of repulsion between them as shown in Figure 5.

"Paired nodes" is a pair of nearby nodes, Q_i and Q_j, where $D_{ij} < R_C$, and Q_j is the nearby node that has shorter distance away from Q_i than all other unpaired nodes.

Q_k is the nearest node away from Q_i, but Q_k has already paired with another node, so it could not become the paired node with Q_i. It means that Q_i must search for an unpaired node which has the shortest distance away from Q_i.

As shown in Figure 6, the repulsion between two nodes should be considered when D_{ij} is shorter than R_C and can be defined as follows:

$$\left\|\vec{F}_{ij}\right\| = \begin{cases} \dfrac{k_R}{{D_{ij}}^2}, & \left(D_{ij} < 2R\right), \\ 0, & (\text{otherwise}), \end{cases} \tag{2}$$

where k_R is the coefficient of repulsion force, which is set as a constant "1" [8].

The repulsion from Q_i to Q_j is $\vec{F}_{ij}$, while $\vec{F}_{ji}$ is the repulsion from Q_j to Q_i. These two forces comply with the principle that $\vec{F}_{ij} = -\vec{F}_{ji}$. The angle between $\vec{F}_{ji}$ and $\vec{C}_i$ is θ_i. Q_i' represents the intersectional point of $\vec{C}_i$ and the clear perceptive circle.

In order to move the force point without changing the direction and the value of force, we suppose that Q_i can be moved to Q_i' along the border of fan-shaped region. This process is equivalent to actually moving $\vec{F}_{ji}$ from Q_i to Q_i' along the direction of $\vec{C}_i$. The force at Q_i' can be decomposed into two components. One is $\vec{F}_{ji_T}$ along the direction of tangent line; the other is $\vec{F}_{ji_C}$ along the direction of normal axis of tangent line. With the influence of $\vec{F}_{ji_T}$, the fan-shaped region gains rotatable trend around Q_i. At the same time, $\vec{C}_i$ is going to be adjusted along the direction of $\vec{F}_{ji_T}$.

The perceptive directions of nodes belong to uniform distribution after initial deployment. Hence, there might be some special situations, for example, when θ_i equals π or zero and $\vec{F}_{ji_T}$ is zero, the perceptive direction need not be adjusted. Nevertheless, when the respective angles θ of two paired nodes are both zero, it would inevitably lead to the worst condition of coverage.

In practice, the value of RSSI (received signal strength indicator) [14, 15] can be used to determine the nearest node away from the current node. The two paired nodes have to interchange messages to inform each other the distance and perceptive direction information. If both θ_i and θ_j are zero, it can be judged as the worst situation of coverage. When θ_i and θ_j are both π, we add some little turbulence, so that the FOV of such node could also be adjusted to further increase the coverage.

3.3. Perceptive Direction Adjustment. Because of the virtual force between two paired nodes, adjustment of angular magnitude turns out to be a trouble in coverage problem. We propose two calculating methods, that is, one is linear-relation-based algorithm and the other is a mechanism-based approximate algorithm.

3.3.1. Linear-Relation-Based Algorithm (LRBA). LRBA is used to calculate the angular magnitude marked by $\Delta\varphi$ that needs to be adjusted. Figure 7 is the paradigm based on LRBA.

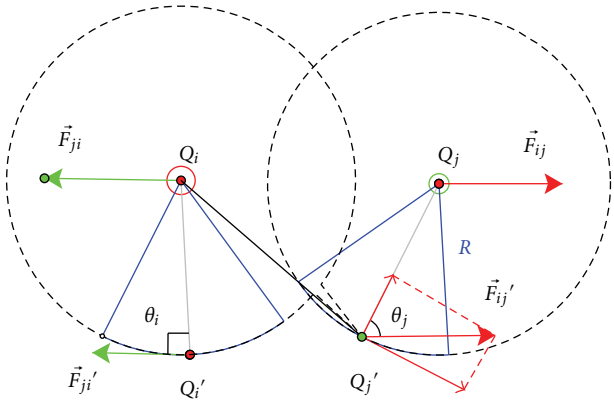


Figure 7: The paradigm based on LRBA.

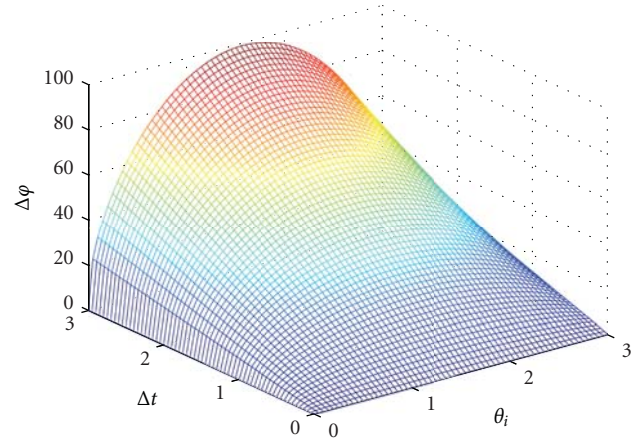


Figure 8: Relationship of θ_i, Δt, and $\Delta\varphi$.

In Figure 7, θ_i is $\pi/2$ and $\vec{F}_{ji_T} = \vec{F}_{ji}$. The FOV of Q_i gains the trend of clockwise rotating around Q_i. It is assumed that the effects can neutralize each other so that $\vec{F}_{ji_T}$ would be zero when $\vec{F}_{ji}$ and $\vec{C}_i$ have the same direction, and it may eventually keep in this state. $\vec{F}_{ji_T}$ generates the rotation of $\vec{C}_i$, whose value is cumulatively changed for $\|\vec{F}_{ji}\|$ during the process of rotating from one state that θ_i is $\pi/2$ to another that θ_i is zero. Meanwhile, $\Delta\varphi$ is adjusted for $\pi/2$. So with the influence of $\vec{F}_{ji} \cdot \sin\theta_i$, $\Delta\varphi$ can be depicted as follows:

$$\frac{\vec{F}_{ji}}{\pi/2} = \frac{\vec{F}_{ji} \cdot \sin\theta_i}{\Delta\varphi}. \tag{3}$$

Equation (3) can generate the following:

$$\Delta\varphi = \frac{\pi}{2} \cdot \sin\theta_i. \tag{4}$$

The range of θ_i is $[0, \pi]$ and $\sin\theta_i$ gains the range of $[0, 1]$. The range of $\Delta\varphi$ that should be adjusted is $[0, \pi/2]$ for each image sensor node in networks. This conclusion fits for actual demand.

3.3.2. Mechanism-Based Approximate Algorithm (MBAA). Because $\vec{F}_{ji_C}$ is changing all the time, it is impossible that the FOV can circle at an immutable speed. Obviously, the larger $\vec{F}_{ji_T}$ is, the faster the rotational speed is. In fact, the change of θ_i from zero to π does not affect the value of $\vec{F}_{ji}$ according to (2). However, D_{ij} has a significant impact on $\vec{F}_{ji}$. In the range of $[0, \pi/2]$, $\vec{F}_{ji_T}$ gradually becomes larger, while in the range of $[\pi/2, \pi]$, it gradually becomes smaller:

$$\left\|\vec{F}_{ji_T}\right\| = \omega^2 R, \tag{5}$$

$$\Delta\varphi = \omega \cdot \Delta t. \tag{6}$$

In (5), R is the clear perceptive radius of image sensor node, and ω is the angular velocity of the rotation of effective sensing area. Δt represents the time for an adjustment. The relationship between $\Delta\varphi$ and ω is described by (6). Equation (7) can be derived from the combination of (2), (5), and (6):

$$\Delta\varphi = k_1 \cdot \Delta t \cdot (\sin\theta_i)^{1/2},$$
$$k_1 = \frac{1}{D_{ij}} \cdot \sqrt{\frac{k_R}{R}}. \tag{7}$$

θ_i and Δt both have influence on the value of $\Delta\varphi$. Their relation is shown in Figure 8.

The range of $\Delta\varphi$ is so wide that it would directly aggravate the energy consumption while adjusting perceptive direction without some constrains. So it will assume that Δt is a fixed value in an adjustment in our proposed methods, regardless of the rotation angle value is very large or very small. And (7) can be simplified into (8).

In (8), $\Delta\varphi$ is affected by D_{ij}, R, and Δt. D_{ij} and R are both constant after initial deployment. So the value of $\Delta\varphi$ is directly decided by Δt. As described above, the time for any one adjustment is equivalent. So it is roughly considered that the impact of microvariables can be ignored, and k_2 can be just normalized without affecting the relationship between $\Delta\varphi$ and θ_i as shown in Figure 9. The range of $(\sin\theta_i)^{1/2}$ is $[0, 1]$, which is feasible for adjusting the perceptive direction of sensor nodes:

$$\Delta\varphi = k_2 \cdot (\sin\theta_i)^{1/2},$$
$$k_2 = \frac{\Delta t}{D_{ij}} \cdot \sqrt{\frac{k_R}{R}}. \tag{8}$$

Equations (4) and (8) are consistent with the needs of practical application and will be elaborated in Section 4.2.

4. Implementation

4.1. Detailed Steps. The FVPTR method proposed in this paper follows steps below with the premise of the network is connected.

(I) deploy N sensors randomly in the monitored region. The sink broadcasts the hop-explored packet. After receiving it, each sensor node records the hop information and adds

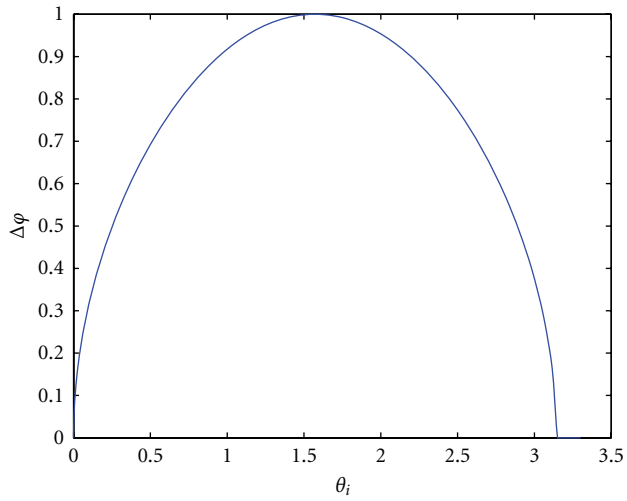


Figure 9: Relationship between θ_i and $\Delta\varphi$.

Begin:
γ is the value of rotation; ε is a given tiny value; β is a value that is larger than ε (we assume that β=10ε); "M" is defined as half of the pixel width value of current image.
A current captured image is scanned by the node.
The node generates a random number marked with *RAND* in the range of $[-1, +1]$.
A. If *itcannot identify any "Video Event"* {
$\gamma = \beta$;
GOTO B;
}
If *a "Video Event" occurs* {
$\gamma = \varepsilon$;
Calculate the current *ACV*;
$\Delta d = |M - ACV|$;
If $\Delta d < \varepsilon$, **GOTO** End;
Else GOTO B;
}
B. If $RAND > 0$
The perceptive direction of node is counterclockwise adjusted for γ;
Else{
Counterclockwise rotates for γ; ***GOTO*** A;
}
End.

Algorithm 1

one to current hop value then rewrites the packet and sends it out. Finally, each sensor node sends packet back to the sink to notify the hop information. Then the sink chooses those sensor nodes, with the largest hop value, records them as boundary nodes and sends "video event" searching command to them.

(II) sensor nodes that accept video event searching command start up video event searching algorithm. It is described as in Algorithm 1.

Table 1: Comparison of different algorithms ($\alpha = \pi/6$, $R = 60$ m, $N = 100$).

	Initial coverage (%)	Adjusted coverage (%)	Coverage enhancement (%)
LRBA	48.67	53.12	4.45
MBAA	48.58	51.88	3.30
Mixed Superposition Algorithm	48.62	54.07	5.45

(III) the nodes that do not accept video event searching command start virtual-potential-based paired repulsion algorithm as in Algorithm 2.

(IV) each nonboundary node adjusts its perceptive direction in terms of $\Delta\varphi$.

4.2. Remark. The result of step I can effectively distinguish boundary nodes from nonboundary nodes. No doubt that it is the foundation and prerequisite for adjusting perceptive direction of two different kinds of nodes.

Step II elaborates the whole process of video event searching algorithm. It is important for initialization, since actual testing should be done to make sure that image sensor nodes can effectively keep sensitive to "video event" such as strong light source without responding to general event.

In step III, if the current node cannot receive confirmation-paired packet from target node within the given time slot τ, it will search for another node nearby in the range of R_C for pairing. When θ_i and θ_j are both π, it is dispensable to adjust nodes' perceptive direction. If the current node cannot search any target node to complete the pairing process, it will expect the subsequent executions.

From the perspective of optimizing overall coverage, two improved algorithms based on (4) and (8) will be further proved in Section 5.

5. Simulation

5.1. Single-Time Adjustment. The value of FOV is $\pi/3$ in line with actual test in Section 2.2, and α is $\pi/6$. N nodes are distributed in a square region of 500 m $*$ 500 m. Their perceptive directions are subject to uniform distribution in the range of $[0, 2\pi]$. Different equations are used to calculate $\Delta\varphi$ of each node. All simulation data listed below are the average value of numerous (100 times) test results.

The results of simulation for LRBA, MBAA, and mixed superposition algorithm are shown in Table 1. The calculation on $\Delta\varphi$ of mixed superposition algorithm can be depicted as follows:

$$\Delta\varphi = \left| \frac{\pi}{2} \cdot \sin\theta_i + (\sin\theta_i)^{1/2} \right|. \quad (9)$$

Obviously the enhancement of mixed superposition algorithm seems to be better than two other methods. However, owed to the stochastic character, mixed superposition algorithm reduces the complexity of computation and the degree of enhancement.

Begin:
a. Each current node searches for other nodes in its range of R_C and chooses one that has the maximum RSSI value (inversely proportional to the distance), then marks it as target node;
b. Request-paired packet is sent from current node to target node.
If *target node is idle* {
It responses the request and marks the sender as its paired node, then sends back confirmation-paired packet to the current node; so the target node becomes the paired node of current node; **GOTO** c;
}
Else {
The target node is in the process of pairing with other nodes; the current node will wait for the time slot of τ then select another target node with the second largest RSSI value and continue pairing process with reference to a.
}
Repeat the above procedure until the current node finds its all neighbors are paired with others or the time domain is over.
c. Information is exchanged between paired nodes.
If *the worst condition of coverage occurs* ($\theta_i = \theta_j = 0$)
{
A *NUMBER* is generated randomly in the range of $[-1, +1]$ by the node sent the request-paired packet.
If $NUMBER > 0$
The perceptive direction is clockwise adjusted for $\pi/6$;
Else counterclockwise rotates for $\pi/6$;
Taking inverse operation to *NUMBER* value, the sensor obtains $\overline{NUMBER}$ and sends it to another paired node. Then, the paired node does the same operation as follows:
If $\overline{NUMBER} > 0$
The perceptive direction is clockwise adjusted for $\pi/6$;
Else counterclockwise rotates for $\pi/6$;
}*//end if*
d. Force analysis between two paired nodes: each node calculates θ, then computes $\Delta\varphi$ according to(4) and (8) respectively;
End.

Algorithm 2

Whatever any algorithm, single-time adjustment fails to gain enhancement more than 6%, which cannot meet the actual demands.

5.2. Multiple-Time Adjustment. After the first successful pairing, each node starts the second pairing. For each node, it neglects the node that has paired with itself in the first time and searches for another node in the range of R_C again and then does the same operations of adjustment. Through repeated tests, the best coverage enhancement occurs after adjusting for three times. In the first time, (4) is used to adjust and (8) is used for the second adjustment. Finally, mixed subtraction (not the mixed superposition) algorithm is applied in the third time, in which the absolute value of (8) subtracted from (4) is used, shown as in the following:

$$\Delta\varphi = \left| \frac{\pi}{2} \cdot \sin\theta_i - (\sin\theta_i)^{1/2} \right|. \tag{10}$$

It is significant for three-time adjustment to use different algorithms. In the second pairing, the current node searches

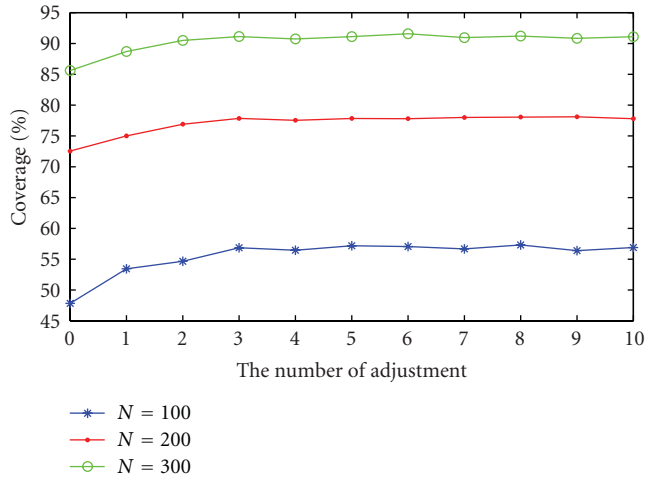


Figure 10: The effect of the number of adjustment.

Table 2: Results of algorithm that adjusts for three times corresponding to Figure 11.

Condition: $\alpha = \pi/6$, $R = 60$ m, $N = 100$	Coverage (%)
Initial deployment	47.84
First adjustment	53.45
Second adjustment	54.67
Third adjustment	56.86
Coverage enhancement	9.02

Table 3: Average results of algorithm that adjusts for five times with other parameter values.

Condition: $\alpha = \pi/4$, $R = 60$ m, $N = 200$	Coverage (%)
Initial deployment	72.13
First adjustment	76.37
Second adjustment	78.87
Third adjustment	81.14
Fourth adjustment	81.72
Fifth adjustment	82.05
Coverage enhancement	9.92

for another node that has the shortest distance from it. Nonetheless, the distance should be longer than that of the first paired node from the current node, as similarly in the third adjustment. According to (2), the farther the distance between two paired nodes is, the less the force between them should be. Moreover, the less the force is, the smaller $\Delta\varphi$ is. So the farther the distance is, the smaller the adjusting angle value is. The relation of the average value of each algorithm's $\Delta\varphi$ can be arranged that LRBA in the first adjustment > MBAA in the second adjustment > mixed subtraction algorithm in the third adjustment.

We examine the effect of the number of adjustment. In the condition that $\alpha = \pi/6$, $R = 60$ m, and N is set 100, 200, and 300, respectively, as shown in Figure 10, with the increase of the number of adjustment, the coverage increases linearly until the number of adjustment reaches 3 and then becomes saturated when the number of adjustment is above 3. So we assume the number of adjustment is 3 in the following simulation.

In the condition that $\alpha = \pi/6$, $R = 60$ m, and $N = 100$ (2α = FOV), Figure 11 shows the simulation process of three-time adjustment.

The boundary nodes are colored with green, while non-boundary nodes are blue. The coverage of networks is enhanced evidently as shown in Figure 11(d). The detailed values are listed in Table 2.

In the premise of initial deployment, coverage is 72.13%, by 100 times tests, Table 3 displays the comparative average results on adjusting for five times with the parameters $\alpha = \pi/4$, $R = 60$ m, and $N = 200$.

From Tables 2 and 3, after three adjustments, the coverage improves nearly twice that of the single-time adjustment. By actual tests, there is only little difference between the results of three-time adjustment and adjusting for more times. However, the unsatisfactory results of declining in coverage occur after adjusting for more than ten times. Because of the random feature of algorithms, the more the times of pairing are, the more difficult to control the final outcome of adjustment is.

5.3. Coverage Enhancing versus Different Parameters

Case 1 (changes in the value of "α"). As shown in Figure 12, the enhancement gains the optimal value when $\alpha = \pi/4$ in all of four algorithms. The worst situation occurs while α is $\pi/3$. The changes generated by diversity of α in different algorithm roughly have the same trend which turns out to be Z-shaped discipline in Figure 12.

Case 2 (changes in the value of "N"). Coverage comparison of different algorithms is expressed in Figure 13 through changing the number of nodes N, while $\alpha = \pi/6$ and $R = 60$ m.

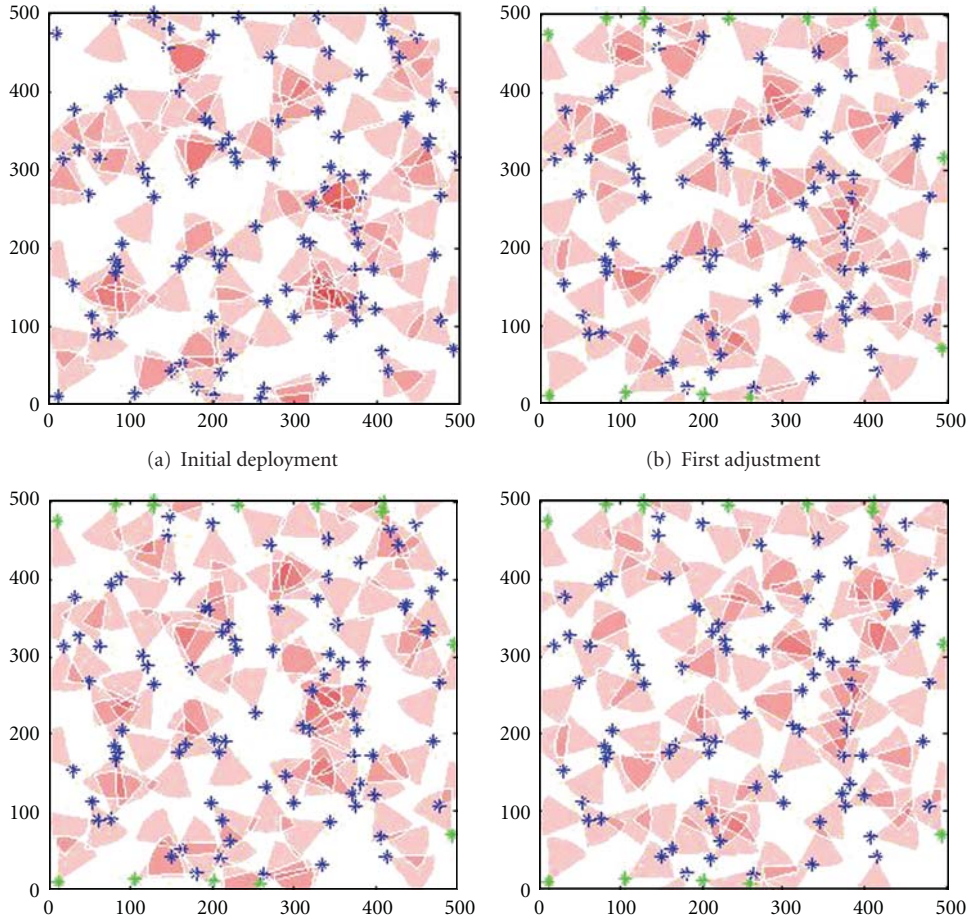


Figure 11: The results of three-time adjustment.

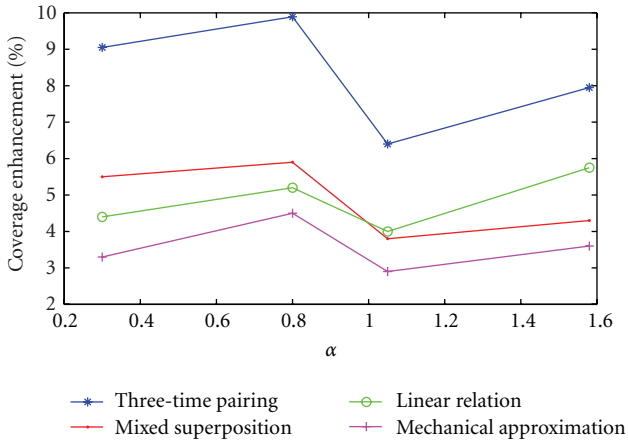


Figure 12: Influence on enhancement caused by α.

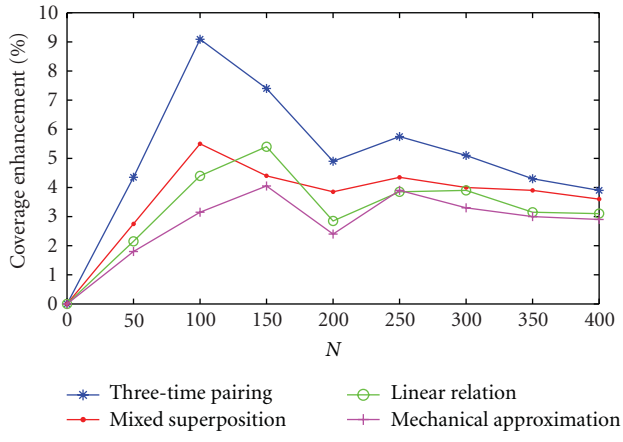


Figure 13: Influence on enhancement caused by N.

As shown in Figure 13, the effect on enhancement caused by different algorithms can be sorted that three-time pairing > mixed superposition > linear ration > mechanical approximation. And the following conclusions can be drawn after analyzing the data in Figure 13.

(a) Different algorithms gain different peak position. The peak of three-time pairing and mixed superposition appear

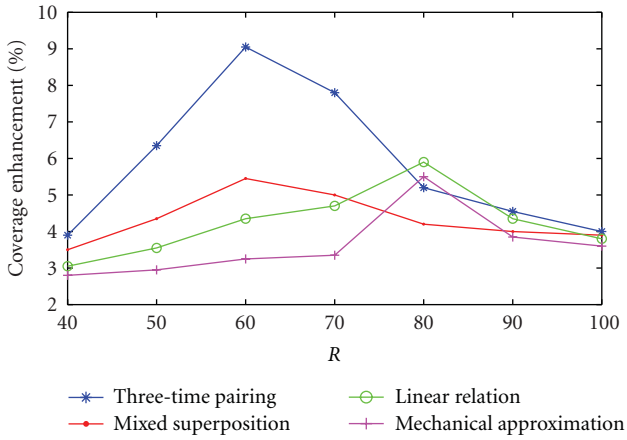


Figure 14: Influence on enhancement caused by R.

in vicinity of 100 nodes while that of two others occur in vicinity of 150 nodes.

(b) In vicinity of 200 nodes, the effect of each algorithm turns out to a certain degree slow-paced.

(c) When the number of nodes is beyond 250, the coverage improvement of each algorithm gradually declines.The reasonable explanations for those phenomena are as follows.

Situation A. The average values of angles that need to be adjusted of three-time pairing and mixed superposition are larger than that of two other algorithms. Meanwhile, once reaching the coverage peak value, required nodes number of three-time pairing and mixed superposition is less than lined relation and mechanical approximation. It is confirmed that there must be some relationship between the peak of enhancement and angles that need to be adjusted.

Situation B. In vicinity of 200 nodes, the NR/NC (quotient of network redundancy dividing network coverage) gains its minimum value. In other words, the network configuration resource is utilized adequately when the number of nodes is about 200. So in this situation, the outcome of adjustment is not remarkable.

Situation C. With the continuous increasing on the number of nodes in a limited area, the coverage enhancement gradually becomes saturated.

Case 3 (changes in the value of "R"). Figure 14 shows the condition that N is 100, α is $\pi/6$, and the value of R is gradually increasing along the horizontal axis direction.

The following conclusions can be drawn.

(A) The peaks of three-time pairing and mixed Superposition occur in vicinity of 60 m while that of linear ration and mechanical approximation do in vicinity of 80 m.

(B) The changes of enhancement of four algorithms keep the same trend, which rises firstly and then falls down after reaching the peak.

(C) Except in vicinity of 80 m and 90 m, the enhancement of algorithms can be sorted that three-time pairing > mixed superposition > linear ration > mechanical approximation.

The trend of changes in Figure 14 can be understood from several aspects.

Situation A. Like the situation described in Figure 13, the peaks of different algorithms emerge at different positions. From the macropoint of view, the impact on coverage enhancement caused by average value of angle adjusted shows that the peak position of linear ration and mechanical approximation lags behind two other algorithms.

Situation B. When the radius of nodes tends to zero, analysis of coverage cannot be set up and coverage enhancement cannot be realized as well. When the radius tends to infinite the complete coverage can be achieved after initial deployment. In practice, the enhancement would be conspicuous only in a certain range of radius.

Situation C. As shown in Figure 14, ignoring the effect of peak value, the three-time pairing algorithm shows its advantage mainly due to taking into account the impacts of other neighboring nodes.

6. Conclusion

In this paper, based on virtual potential field, the paired tangent point repulsion for nonboundary sensor nodes and fuzzy image recognition for boundary sensor nodes realize the enhancement of perspective coverage, together with LRBA, MBAA, and mixed superposition algorithm for rotation angle adjustment. Furthermore, through simulation experiments based on the above algorithms, three-time adjustment method gains better performance than single one and more satisfactory cost-efficacy ratio than more-time adjustment method.

However, there exists defects in the algorithm execution of FVPTR, for example, some nodes could not find their respective paired partner such as those with only one neighbor or the remaining one alone from the odd numbers of total. This is the emphasis in the further research, and meanwhile coverage issue on video or audio sensors networks will be developed in the next work.

Acknowledgments

The subject is sponsored by the National Natural Science Foundation of China (nos. 60973139, 61003039, 61170065, and 61171053), Scientific and Technological Support Project (Industry) of Jiangsu Province (nos. BE2010197, BE2010198), Natural Science Key Fund for Colleges and Universities in Jiangsu Province (11KJA520001), the Natural Science Foundation for Higher Education Institutions of Jiangsu Province (10KJB520013, 10KJB520014), Academical Scientific Research Industrialization Promoting Project (JH2010-14), Fund of Jiangsu Provincial Key Laboratory for Computer Information Processing Technology (KJS1022), Postdoctoral Foundation (1101011B), Science and Technology Innovation Fund for Higher Education Institutions

of Jiangsu Province (CXZZ11-0409), the Six Kinds of Top Talent of Jiangsu Province (2008118), Doctoral Fund of Ministry of Education of China (20103223120007, 20113223110002), and the Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (yx002001).

References

[1] M C. Zhao, J. Y. Lei, M. Y. Wu et al., "Surface coverage in wireless sensor networks," in *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM '09)*, pp. 109–117, Rio de Janeiro, Brazil, 2099.

[2] H. D. Ma, X. Zhang, and A. L. Ming, "A coverage-enhancing method for 3D directional sensor networks," in *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM '09)*, pp. 2091–2795, Rio de Janeiro, Brazil, 2009.

[3] X. Bai, S. Kumar, D. Xuan, Z. Yun, and T. H. Lai, "Deploying wireless sensors to achieve both coverage and connectivity," in *Proceedings of the 7th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MOBIHOC '06)*, pp. 131–142, Florence, Italy, May 2006.

[4] X. Han, X. Cao, E. L. Lloyd, and C. C. Shen, "Deploying directional sensor networks with guaranteed connectivity and coverage," in *Proceedings of the 5th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks, (SECON '08)*, pp. 153–160, San Francisco, Calif, USA, June 2008.

[5] A. Jing and A. A. Alhussein, "Coverage by directional sensors in randomly deployed wireless sensor networks," *Wireless Network Applications*, vol. 11, no. 1, pp. 21–41, 2006.

[6] H. Mohamed and A. Hossien, "A probabilistic coverage respective for wireless sensor networks," in *Proceedings of the 15th IEEE International Conference on Network Protocols (ICNP '07)*, pp. 41–50, Beijing, China, 2007.

[7] G. Wang, G. Cao, and T. F. La Porta, "Movement-assisted sensor deployment," *IEEE Transactions on Mobile Computing*, vol. 5, no. 6, pp. 640–652, 2006.

[8] D. Tao, H. D. Ma, and L. Liu, "Virtual potential field based coverage-enhancing algorithm for directional sensor networks," *Journal of Software*, vol. 18, no. 5, pp. 1152–1163, 2007.

[9] Y. Zou and C. Krishnendu, "Sensor deployment and target localization based on virtual forces," in *Proceedings of the 22nd Annual Joint Conference of the IEEE Computer and Communications*, vol. 2, pp. 1293–1303, Francisco, Calif, USA, 2003.

[10] H Oztarak, A Yazici, D Aksoy et al., "Multimedia Processing in Wireless Sensor Networks," in *Proceedings of the 4th International Conference on Innovations in Information Technology (IIT'07)*, pp. 78–82, Dubai, UEA, 2007.

[11] S. Stanislava and B. H. Wendi, "On the coverage problem in video-based wireless sensor networks," in *Proceedings of the 2nd International Conference on Broadband Networks (BROADNETS '05)*, vol. 2, pp. 932–939, Boston, Mass, USA, October 2005.

[12] A Howard, M. J. Matari'c, and G. S. Sukhatme, "Mobile sensor network deployment using potential field: a distributed scalable solution to the area coverage problem," in *Proceedings of the 6th International Symposium on Distributed Autonomous Robotics Systems (DARS'02)*, pp. 299–308, Japan Scientific & Technical Information, Fukuoka, Japan, 2002.

[13] P. Jiao, H. J. Wang, and F. G. Ding, "Local path planning method for autonomous underwater vehicle based on virtual field force," *Ship Building of China*, vol. 48, no. 1, pp. 76–81, 2007.

[14] K Benkic, M Malajner, P. Planinsic et al., "Using RSSI value for distance estimation in Wireless sensor networks based on ZigBee," in *Proceedings of the 15th International Conference on Systems, Signals and Image Processing (IWSSIP '08)*, pp. 303–306, Bratislava, Czechoslovakia, 2008.

[15] R. H. Wu, Y. H. Lee, H. W. Tseng et al., "Study of characteristics of RSSI signal," in *Proceedings of the IEEE International Conference on Industrial Technology, (ICIT'08)*, pp. 1–3, Chengdu, China, 2008.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 603709, 6 pages
doi:10.1155/2012/603709

Research Article

An Energy Supply System for Wireless Sensor Network Nodes

Guoqiang Zheng,[1] Liwei Zhang,[1] and Jishun Li[2]

[1] *School of Electronic Information Engineering, Henan University of Science and Technology, Room 48, Xiyuan Road, Jianxi District, Luoyang 471003, China*
[2] *Henan Key Laboratory for Machinery Design and Transmission System, Henan University of Science and Technology, Luoyang 471003, China*

Correspondence should be addressed to Guoqiang Zheng, xinzhzhangliwei@126.com

Received 26 November 2011; Revised 25 March 2012; Accepted 25 March 2012

Academic Editor: Sabah Mohammed

The power source is a critical obstacle for wireless sensor network nodes. In order to prolong the lifetime of wireless sensor networks, this paper presents an energy supply system that uses a specially designed broadband piezoelectric energy harvesting technology for sustaining the operation of wireless sensor network nodes. The proposed energy supply circuit can apply an optimal control voltage to the piezoelectric element to ensure impedance matching between the vibration source and the energy supply system in nonresonance frequencies. As compared to the conventional piezoelectric energy harvesting circuit, it is shown that the efficiency has been increased 4 times. In this work, the overall system structure, the function modules design, and the performance testing analysis are illuminated in detail. Experimental results reveal that this energy supply system can significantly improve power within the wide bands by the active piezoelectric energy harvesting technology and enable wireless sensor network nodes to operate normally.

1. Introduction

Wireless sensor networks have been of great interests over the last few decades. Wireless sensor networks are the integration of sensor technology, embedded computing technology, modern network and wireless communication technology, distributed information processing technology, and so on. They can be used to monitor, sense, and collect the information on the environment or objects by microsensors and transmit these information to the users. Therefore, they have gained numerous applications such as military defense, industry and agriculture, city management, biological and medical treatment, and environmental monitoring [1–3].

However, wireless sensor networks are not rapid commercialization as people have expected. One of the most critical bottlenecks is the energy supply problem for wireless sensor network nodes. At present, the wireless sensor network node generally uses traditional chemical battery. Because of the large numbers of devices and their small size, changing the battery is unpractical or simply not feasible. It cannot fully meet the development requirements of wireless sensor networks. Therefore, more and more attention has been attracted to harvest energy from the surrounding environment to achieve the self-power for wireless sensor networks [4–6]. The combination of an energy harvester with a small-sized rechargeable battery (or with another energy storage system like a thin-film rechargeable battery or a supercapacitor) is the best approach to enable energy autonomy of the network over the entire lifetime. Some possible ambient energy sources are, for instance, photonic energy, thermal energy, or mechanical energy. Because of fast response, low cost, simple structure, no electromagnetic interference, easy manufacture, and so forth, piezoelectric energy harvesters are suitable for wireless sensor networks [7].

Piezoelectric energy harvesters convert mechanical strain energy into electricity. When the resonance frequency of the harvester matches with the input frequency of vibration, the harvester can output the maximum electric power. In practical applications, there are two important problems in piezoelectric energy harvesters. Firstly, the bandwidth of the piezoelectric device is relatively narrow. The vibration status is often unsteady and varies from applications to applications

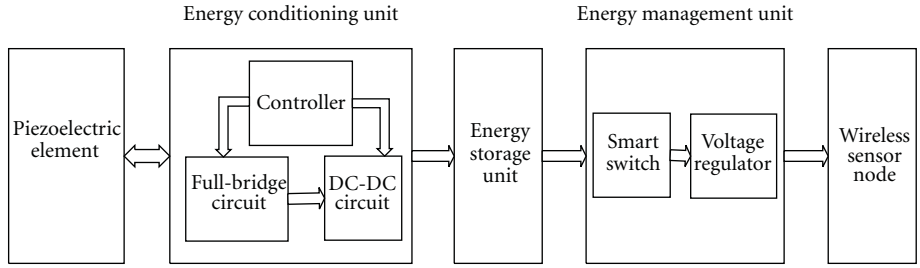


Figure 1: Schematic of the energy supply system.

and from time to time. If the vibration frequency is slightly away from the resonance frequency, the output electricity will significantly decrease. To enhance the harvested power, we need to track the vibration status dynamically and adjust the piezoelectric energy harvester for best load matching. Secondly, the energy harvesting efficiency is small, and the harvested energy is very little.

The purpose of this paper is placed on resolving the two aforementioned challenges by using the designed ultra-low-power energy supply circuit to harvest maximal energy for wireless sensor network nodes. The energy supply system can run an adaptive active piezoelectric harvesting technology to generate an optimal control voltage and improve harvested energy in the broadband. The remainder of this paper is organized as follows. The related work in the literature is presented in Section 2. Section 3 gives an overall design scheme of the energy supply system. Section 4 details the energy conversion unit and the active piezoelectric energy harvesting technology. The energy management unit is given in Section 5. Section 6 gives the system simulation, results, and discussion. Finally, we draw the conclusions of this paper in Section 7.

2. Related Work

A great amount of researches has been conducted about the energy harvesting technology as a self-power source for wireless sensor network nodes. Shwe et al. exploited a temperature sensing system which harvested the power energy from surrounding machinery vibration by the piezoelectric generator [8]. The approach is very useful in real-time remote monitoring of the machine temperature by sensor network in industries. Tan et al. presented an energy harvesting circuit design of piezoelectric pushbutton generator for wireless radiofrequency (RF) transmitter [9]. When the piezoelectric push button is depressed, 67.61 mW of electrical energy is scavenged and it is sufficient to transmit 12-bit digital word information.

As described above, the harvested energy by the piezoelectric energy harvester is very small. At the same time, it is greatly affected by the vibration status (i.e., magnitude and frequency). Therefore, previous research works have addressed the issue of maximizing the harvested electrical power. Several optimization schemes have been proposed in the literature [10, 11]. They are based on one of these principles: reducing the power loss in rectifying diodes [12], improving the extracted power by using an adaptive circuit [13, 14], adaptive control of the rectified DC voltage [15], or adjusting the natural frequency of piezoelectric energy harvesters [16]. At present, there are three predominant energy harvesting circuits: the passive diode-rectifier circuit [17], the semiactive circuit [18–20], and the active circuit [21]. The passive diode-rectifier circuit is the simplest technology, and its efficiency is the lowest. In the semiactive circuit, the output voltage can be processed nonlinearly by switched control circuit to increase its magnitude and change its phase so that the harvested electrical energy is maximized. In the active circuit, appropriate electrical boundary conditions applied to the piezoelectric element can push the harvested energy to the limits of the piezoelectric harvester. To improve the adaptability of the energy supply system, we propose an adaptive active piezoelectric technology in this paper with the natural frequency of the energy supply system easily adjusted by changing the amplitude and phase of its control voltage.

3. Energy Supply System

The energy supply system that can be used to convert the energy of ambient mechanical vibrations to electricity is used to power the wireless sensor node. In order to increase the generated power and convert more mechanical energy effectively, an energy supply system must be employed. Figure 1 presents a schematic of the proposed system. It contains a piezoelectric element, an energy conditioning unit, an energy storage unit, and an energy management unit.

The piezoelectric element converts the external vibration mechanical energy to alternating power and outputs electrical energy to the energy storage unit through the energy conditioning unit. In the energy conditioning unit, the controller runs an active piezoelectric energy harvesting technology which will be discussed in the next chapter. It outputs an optimal control voltage applied to the full-bridge circuit and the DC-DC circuit. The energy storage unit which is used to store the generated electrical energy is commonly a rechargeable battery or a supercapacitor. The energy management unit contains two parts, a smart switch and a voltage regulator, and monitors the voltage of the energy storage unit. When the voltage of the energy storage unit is in the setting range, the energy management unit can output a constant voltage to power the wireless sensor node.

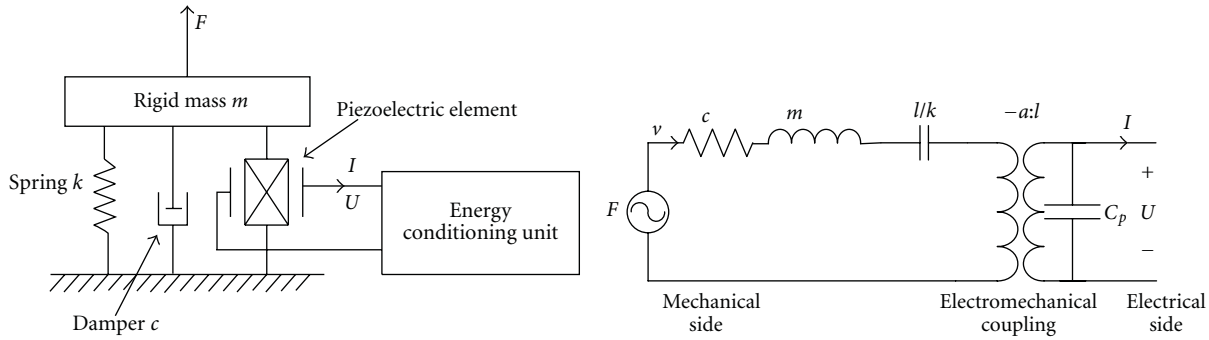


Figure 2: Electromechanical coupling model and equivalent circuit model.

4. Energy Conditioning Unit

The energy conditioning unit is the most important part of this energy supply system. It determines the harvesting power of the energy supply system and affects the lifetime of the wireless sensor node. An active piezoelectric energy harvesting technology is used to this energy conditioning unit. When the frequency of the external exciting force changes, it can adaptively generate an optimal control voltage applied to the piezoelectric element. Therefore, this system can be in resonant state and realizes impedance matching. Figure 2 presents the electromechanical coupling model and the equivalent circuit model of the piezoelectric energy harvester [22].

In this basic approach, v is the exciting velocity, m is the rigid mass, c represents the damping coefficient, the spring k corresponds to the stiffness of the mechanical structure, α is the force factor, and C_P corresponds to the plate capacitance of piezoelectric element. U is the output voltage, and I is the output current of the piezoelectric element.

Assuming that the external exciting force F is sinusoidal, the impedance of the mechanical side is given by

$$\vec{Z}_m = c + j\left(\omega m - \frac{k}{\omega}\right), \tag{1}$$

where $\omega = 2\pi/v$. When the impedance of the electrical part corresponds to the mechanical part, the maximum energy can be harvested. The optimal matching impedance of the mechanical side is

$$\vec{Z}_m = c + j\left(\frac{k}{\omega} - \omega m\right). \tag{2}$$

Therefore, the optimal impedance of the electrical part can be expressed as

$$\vec{Z}_m = \frac{\omega^2 m - k + j\omega c}{\omega^2 c_p - j\omega\left(\omega^2 c_p m - ck\right)}. \tag{3}$$

Starting from (3), the optimal impedance is difficult to be achieved. We propose an active piezoelectric harvesting technology and apply a control voltage to piezoelectric element for generating the equivalent impedance.

The optimal control voltage can be found out

$$\vec{U}_{\text{opt}} = \frac{-\omega c - j(k - \omega^2 m)}{2\alpha\omega c}\,\vec{F} \tag{4}$$

The optimal magnitude of the control voltage can be expressed by

$$U_{\text{mag}} = \sqrt{\omega^2 c^2 + (k - \omega^2 m)^2}\,\frac{F_m}{2\alpha\omega c}. \tag{5}$$

The optimal phase angle between the control voltage and the excitation force can be written as

$$\theta = 180 - \arctan\left(\frac{k - \omega^2 m}{\omega c}\right). \tag{6}$$

To achieve the active piezoelectric energy harvesting technology, a circuit of the energy conditioning unit is shown in Figure 3. When in the external environment vibration occurs, piezoelectric element converts mechanical energy into electrical energy. Because the output energy of the piezoelectric element is alternating, a full-bridge rectifier is essential to transform the AC to the DC. A step-down converter is mainly used to regulate the rectified voltage, so the system can output the maximum power by setting the duty cycle of the DC-DC converter. The power detector detects the power which flows into a constant voltage rechargeable battery and inputs the value of the power to the MSP430F169 controller. The MSP430F169 controller can give the optimal amplitude and phase angle of a control voltage. If we use the excitation force as the reference, the phase generator will generate a square-wave drive signal of the optimal phase angle between the control voltage and the excitation force to control the four MOSFETs of the full-bridge converter which can apply the optimal control voltage to the piezoelectric element.

The full-bridge rectifier adopts a diode rectifier and four N-channel ZVN3320 MOSFETs which are controlled by the phase angle between the control voltage and the vibration force. Firstly, it can convert AC power to the DC power. Secondly, it can apply the optimal control voltage to the piezoelectric element.

Because the power levels associated with the piezoelectric energy harvester are very low in the practical application, a

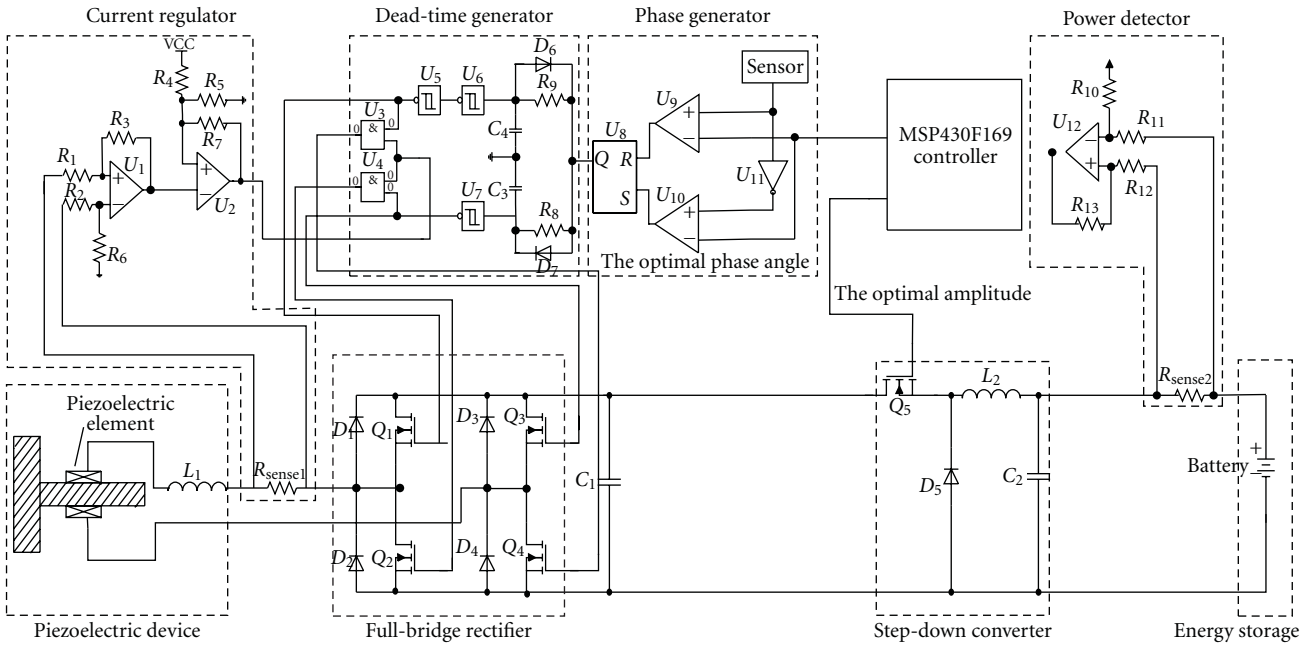


Figure 3: Energy conditioning unit circuit.

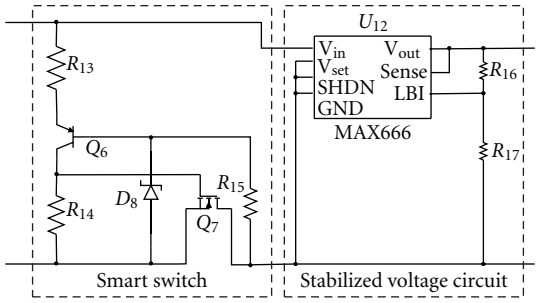


Figure 4: Circuit of the energy management unit.

step-down converter is designed to adjust the rectifier voltage for maximizing the output power. In the implementation, the IRF7853 was chosen as the transistor for its high speed and low power consumption. The optimal magnitude of the control voltage is used to control the transistor to get the optimal rectifier voltage.

In the phase generator, a TS862 by ST is chosen as the comparator for its ultra-low-power consumption. It compares the vibration force signal detected by the accelerometer to the optimal control voltage and generates two pulse signals that are 180° out of phase. Finally, RS trigger transforms the pulse signals to a square-wave phase signal.

To avoid the simultaneous turn-on of the two switches in the same bridge arm, a dead-time generator is essential. SN74LVC3G14 and SN74AHCT08 by TI are chosen as the Schmitt-trigger inverter and the AND gate.

During the voltage transition periods of the piezoelectric element, there may be a large charging/discharging current resulting in large I^2R loss, so a current regulator is developed. It can monitor the output current and output pulse signals to reduce the power loss. At the same time, a power detector is necessary and can be simply built by a differential amplifier with a sense resistance to measure the output power.

5. Energy Management Unit

The energy management unit detects the voltage of the energy storage unit and controls the output power to drive the operation of wireless sense nodes. It contains two parts: a smart switch and a stabilized voltage circuit. Figure 4 presents the circuit of the energy management unit.

When the external vibration occurs, the rechargeable battery of the energy storage unit is charged. If the voltage of the rechargeable battery increases to the threshold voltage which is determined by the voltage stabilizing diode D_8, the transistor Q_6 turns on. At the same time, the resistor R_{14} divides the voltage of the rechargeable battery and makes the MOSFET Q_7 conduct. The rechargeable battery begins to discharge, and the power is transferred to the stabilized voltage circuit. When the voltage of the resistor R_{14} is too low to make the MOSFET conduct, the energy management unit disconnects. The rechargeable battery continues to be charged.

6. Experimental Results and Discussions

Experimental results are taken to validate the energy supply system presented in this paper and to demonstrate the operation of the energy supply system. A sinusoidal exciting force generated by the shaker is applied to the piezoelectric element. First of all, we test the performance of the energy conditioning unit. Under the active piezoelectric energy harvesting technology, the voltage and current of the piezoelectric element are shown in Figure 5. It can be seen that when the vibration force happensat a maximum or minimum value, the voltage of the piezoelectric element will turn in the short time, and the current will be well regulated. Because of the current regulator, the out current is very small among −10 mA to 10 mA.

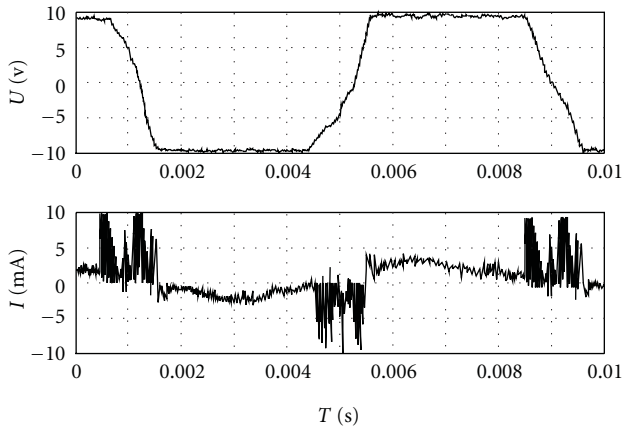


Figure 5: Voltage and current of the piezoelectric element.

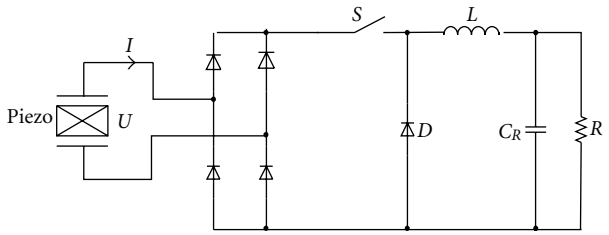


Figure 6: Classic energy harvesting technology.

Ottman et al. design a classic energy harvesting technology for wireless remote power supply, as shown in Figure 6 [23]. It consists of an AC–DC rectifier with an output capacitor, an electrochemical battery, and a switch-mode DC–DC converter that controls the energy flow into the battery.

Compared to the classic energy harvesting technology, Figure 7 presents the experimental and theoretical power of the energy conversion unit. Because of the circuit efficiency, the harvested power is no longer a constant for different excitation frequencies. The maximum power harvested by the active circuit is 9.8 mW at the resonance frequency 85 Hz. In the nonresonant frequencies, the experimental power regulated by the active piezoelectric energy harvesting technology does not quickly reduce. The output power is up to 4 times larger than the power by the classic energy harvesting technology in nonresonance frequencies.

Assuming the input of the energy management unit is a sinusoidal wave, the output voltage of the smart switch is presented in Figure 8. When the input voltage is lower than 3.6 V, the output voltage is 0. When the input voltage is higher than 3.6 V, the output voltage equals to the input value and the smart switch is equivalent to a voltage follower.

7. Conclusions

In this paper, a new energy supply system based on the mechanical vibrations for the wireless sensor node is discussed. An active piezoelectric energy harvesting technology is proposed to make the impedance of the electromechanical coupling system matching. Therefore, the energy supply system can improve the harvested power within the wide bands. At last, the energy supply system can output the constant voltage through the energy management unit to power the wireless sensor node.

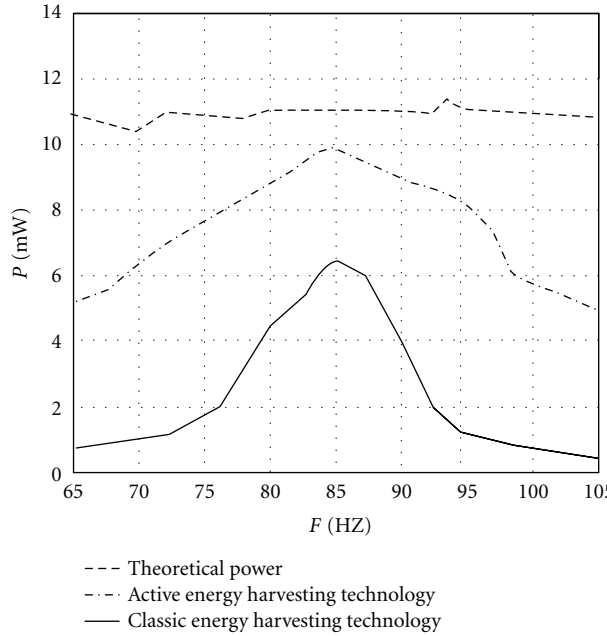


Figure 7: Experimental and theoretical power of the energy conditioning unit.

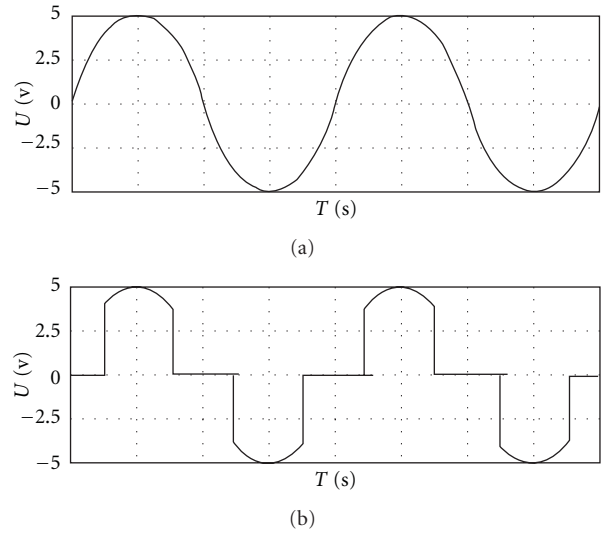


Figure 8: Input and output voltage of the smart switch.

Acknowledgments

This research is supported by the National Hi-Tech Research Development Program (863 Program) (no. 2009AA044902), the Key Scientific and Technological Project of Henan Province (91100210300), the Henan University of Science and Technology Scientific Research Start-up Foundation for

Dr. (no. 09001550), the National Defense Project (JPPT-ZCGX1-1), the Major Project of Basic and Advanced Technology Research of Henan Province (No. 112300413205), the Henan University of Science and Technology Innovation Fund Project for Postgraduates (CXJJ-YJS-Z011), and the Education Department of Henan Province Key Science and Technology Research Project (12B510010).

References

[1] E. P. James, M. J. Tubor, S. P. Beeby et al., "An investigation of self-powered systems for condition monitoring applications," *Sensors and Actuators, A*, vol. 110, no. 1–3, pp. 171–176, 2004.

[2] Z. Wen, Z. Wen, X. He, H. Liao, and H. Liu, "Vibration-based piezoelectric generator and its application in the wireless sensor node," *Chinese Journal of Mechanical Engineering*, vol. 44, no. 11, pp. 75–79, 2008.

[3] R. Torah, P. Glynne-Jones, M. Tudor, T. O'Donnell, S. Roy, and S. Beeby, "Self-powered autonomous wireless sensor node using vibration energy harvesting," *Measurement Science and Technology*, vol. 19, no. 12, Article ID 125202, 2008.

[4] P. Li, C. Olmi, and G. Song, "Energy efficient wireless sensor network for structural health monitoring using distributed embedded piezoelectric transducers," in *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems*, vol. 7647 of *Proceedings of SPIE*, March 2010.

[5] J. Y. Lee, J. G. Oh, K. S. Kim, and B. Choi, "The energy conversion system with piezoelectric effect for wireless sensor network," in *Proceedings of the IEEE International Conference on Sustainable Energy Technologies (ICSET '08)*, pp. 820–824, November 2008.

[6] X. Du and H. Yu, "Piezoelectric battery design to harvest ambient vibration energy for wireless sensor nodes," *Applied Mechanics and Materials*, vol. 26–28, pp. 1088–1092, 2010.

[7] J. Bing, L. Ping, C. Lijuan, L. Xianglin, F. Chi, and X. Xiaohui, "A step-up vibration energy harvester based on piezoelectric bimorph cantilever," *Advanced Materials Research*, vol. 347–353, pp. 3068–3076, 2011.

[8] Y. W. Shwe and Y. C. Liang, "Smart dust sensor network with piezoelectric energy harvesting," in *Proceedings of the 6th International Conference on Information Technology and Applications (ICITA '09)*, pp. 184–187, November 2009.

[9] Y. K. Tan, K. Y. Hoe, and S. K. Panda, "Energy harvesting using piezoelectric igniter for self-powered radio frequency (RF) wireless sensors," in *Proceedings of the IEEE International Conference on Industrial Technology (ICIT '06)*, pp. 1711–1716, December 2006.

[10] J. Hannu, J. Juuti, and H. Jantunen, "Recent patents on piezoelectric energy harvester transducer structures," *Recent Patents on Electrical Engineering*, vol. 3, no. 1, pp. 19–24, 2010.

[11] K. Reilly Elizabeth, B. Fred, F. Romy, and W. Paul, "Powering a wireless sensor node with a vibration-driven piezoelectric energy harvester," *Smart Materials and Structures*, vol. 20, no. 12, Article ID 125006, 2011.

[12] E. Lefeuvre, D. Audigier, C. Richard, and D. Guyomar, "Buck-boost converter for sensorless power optimization of piezoelectric energy harvester," *IEEE Transactions on Power Electronics*, vol. 22, no. 5, pp. 2018–2025, 2007.

[13] W. Hongyan, S. Xiaobiao, and X. Tao, "Complex impedance matching for power improvement of a circular piezoelectric energy harvester," *Applied Mechanics and Materials*, vol. 148-149, pp. 169–172, 2012.

[14] G. K. Ottman, H. F. Hofmann, and G. A. Lesieutre, "Optimized piezoelectric energy harvesting circuit using step-down converter in discontinuous conduction mode," *IEEE Transactions on Power Electronics*, vol. 18, no. 2, pp. 696–703, 2003.

[15] L. Mickal, W. Yi-Chieh, and G. Daniel, "Switching delay effects on nonlinear piezoelectric energy harvesting techniques," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 1, pp. 464–472, 2012.

[16] J. Liang and W. H. Liao, "Energy flow in piezoelectric energy harvesting systems," *Smart Materials and Structures*, vol. 20, no. 1, Article ID 015005, 2011.

[17] I. C. Lien and Y. C. Shu, "Array of piezoelectric energy harvesters," in *Energy Harvesting and Scavenging I: Piezoelectric and Synchronization*, vol. 7977 of *Proceedings of SPIE*, Active and Passive Smart Structures and Integrated Systems, 2011.

[18] E. Lefeuvre, A. Badel, A. Benayad, L. Lebrun, C. Richard, and D. Guyomar, "A comparison between several approaches of piezoelectric energy harvesting," *Journa of Physics IV France*, vol. 128, pp. 177–186, 2005.

[19] R. D'Hulst, T. Sterken, R. Puers, G. Deconinck, and J. Driesen, "Power processing circuits for piezoelectric vibration-based energy harvesters," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 12, pp. 4170–4177, 2010.

[20] L. Garbuio, M. Lallart, D. Guyomar, C. Richard, and D. Audigier, "Mechanical energy harvester with ultralow threshold rectification based on SSHI nonlinear technique," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 4, pp. 1048–1056, 2009.

[21] Y. Liu, G. Tian, Y. Wang, J. Lin, Q. Zhang, and H. F. Hofmann, "Active piezoelectric energy harvesting: general principle and experimental demonstration," *Journal of Intelligent Material Systems and Structures*, vol. 20, no. 5, pp. 575–585, 2009.

[22] L. C. J. Blystad, E. Halvorsen, and S. Husa, "Piezoelectric MEMS energy harvesting systems driven by harmonic and random vibrations," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 57, no. 4, pp. 908–919, 2010.

[23] G. K. Ottman, H. F. Hofmann, A. C. Bhatt, and G. A. Lesieutre, "Adaptive piezoelectric energy harvesting circuit for wireless remote power supply," *IEEE Transactions on Power Electronics*, vol. 17, no. 5, pp. 669–676, 2002.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 942080, 9 pages
doi:10.1155/2012/942080

Research Article

Improved Virtual Potential Field Algorithm Based on Probability Model in Three-Dimensional Directional Sensor Networks

Junjie Huang,[1] Lijuan Sun,[2] Ruchuan Wang,[2] and Haiping Huang[2]

[1] *College of Internet of Things, Nanjing University of Posts and Telecommunications, Jiangsu High Technology Research Key Laboratory for Wireless Sensor Networks, and Key Lab of Broadband Wireless Communication and Sensor Network Technology, Ministry of Education, Nanjing 210003, China*
[2] *College of Computer, Nanjing University of Posts and Telecommunications, Jiangsu High Technology Research Key Laboratory for Wireless Sensor Networks, and Key Lab of Broadband Wireless Communication and Sensor Network Technology, Ministry of Education, Nanjing 210003, China*

Correspondence should be addressed to Junjie Huang, 117579663@qq.com

Received 2 November 2011; Revised 12 March 2012; Accepted 13 March 2012

Academic Editor: Sabah Mohammed

In conventional directional sensor networks, coverage control for each sensor is based on a 2D directional sensing model. However, 2D directional sensing model failed to accurately characterize the actual application scene of image/video sensor networks. To remedy this deficiency, we propose a 3D directional sensor coverage-control model with tunable orientations. Besides, a novel criterion for judgment is proposed in view of the irrationality that traditional virtual potential field algorithms brought about on the criterion for the generation of virtual force. Furthermore, cross-set test is used to determine whether the sensory region has any overlap and coverage impact factor is introduced to reduce profitless rotation from coverage optimization, thereby the energy cost of nodes was restrained and the performance of the algorithm was improved. The extensive simulations results demonstrate the effectiveness of our proposed 3D sensing model and IPA3D (improved virtual potential field based algorithm in three-dimensional directional sensor networks).

1. Introduction

The sensory ability of WSNs to physical world is embodied in coverage which is often used to describe the monitoring standard of Quality of Service (QoS) [1, 2]. Coverage optimization in sensor networks plays a significant role in allocating network space, realizing context awareness and information acquisition, and enhancing the viability of networks [3].

Early studies on coverage optimization were based on two-dimensional sensory domain model [4, 5], or 0-1 probability sensory model. For instance, the virtual force algorithm (VFA) proposed by Zou and Chakrabarty [6] moves nodes after all nodes' moving paths have been determined. The authors of [7] proposed an approximate centralized greed algorithm to solve the maximum coverage problem with minimum sensors. Coverage in terms of the number of targets to be covered is maximized, whereas the number of sensors to be activated is minimized. The target involved virtual force algorithm (TIVFA) is proposed by Li et al. [8] and potential field-based coverage-enhancing algorithm (PFCEA) aiming at directional sensor model is proposed by Tao et al. [9]. The authors of [10] discussed multiple directional cover sets problem of organizing the directions of sensors into a group of nondisjoint cover sets in each of which the directions cover all the targets so as to maximize the network lifetime. In [11], which is written by me, I improved the criterion for judging the generation of virtual potential field via cross-set test. However, probability-based three-dimensional sensor networks model is more tending to be in conformity with practical application, for example, the recently arisen multimedia sensor networks [3, 12–16] and underwater networks [17]. In view of three-dimensional sensory model and probability sensory model, overseas and domestic researchers have yielded some research achievements on coverage algorithms in recent years. I proposed

a virtual force algorithm which is applicable with three-dimensional omnidirectional sensory model in [18]. Authors of [19] put forward a three-dimensional directional sensory model and optimized coverage performance using virtual potential field and simulated annealing algorithm. Reference [20] proposed a coverage configuration algorithm based on probability detection model (CCAP). Reference [21] proposed a coverage preservation protocol based on probability detection model (CPP) that makes working nodes in sensor networks as few as possible when network coverage is guaranteed. However, this protocol configures network using centred control algorithm which limits the network scale, and at present most of the literatures have not introduced this probability coverage model into three-dimensional sensor networks. In fact, most practical applied wireless sensor networks are deposited in three-dimensional sensor networks so that it will be more accurate if it is simulated in a three-dimensional space [22, 23]. In [22], Bai et al. proposed and designed a series of connected coverage model in three-dimensional wireless sensor networks with low connectivity and full coverage. In [23], Alam and Haas studied truncated octahedron deployment strategy to monitor network coverage situation. But in most studies, the criterion for the generation of repulsion between two points is simply defined as the situation that the distance between nodes is less than twice the sensory radius, which has been proved inappropriate to directional sensor networks in [11] by me. Therefore, first we analyze the sensory ability of three-dimensional directional sensor based on probability model to design a novel direction-steerable three-dimensional directional sensor model and make use of virtual potential algorithm to adjust node direction to improve coverage effect. In particular, in this paper we proposed a more rational criterion for generation of repulsion in three-dimensional directional sensor networks and introduced a factor called coverage impact factor to estimate the impact on network coverage from the change of sensing direction in advance, to reduce profitless adjustment of sensing direction, save node energy, enhance algorithm performance, and optimize coverage effect.

This paper is organized as follows: Section 2 gives the problem description and related definition. Section 3 describes the improved virtual potential field-based algorithm in three-dimensional directional sensor networks based on probability model. Section 4 describes in detail the algorithm flow. Section 5 verifies the validity of the algorithm via simulation experiments and makes contrast. Section 6 draws conclusion.

2. Coverage Enhancement Issues of Three Directional Sensor Networks

2.1. Problem Formulation. The coverage enhancement issue of three directional sensor networks that is constituted by direction-steerable nodes can be described as follow: how to enhance the degree of coverage by changing the sensing direction so that degree of coverage in target area approaches maximum in condition that a certain number of nodes are randomly distributed in a given three-dimensional target area and part of the area is not covered by nodes and the number and position of nodes are stable.

2.2. Analysis and Definitions on Coverage Enhancement Issue in Directional Sensor Networks. For the purpose of later research, we give the consumptions beforehand, shown as follows.

(1) Every node works independently, namely, sensory task of each node does not depend on others'.

(2) All nodes are isomorphic, namely, all the maximum sensory distance R_S, sensory deviation angle α, and communication radius R_C are equal, respectively, and the communication radius is no less than twice of the maximum sensory distance.

(3) Every node can get the information of its location and sensing direction and the direction is steerable.

Limited by the angle of view, the sensory area of the directional sensor model is abstracted to a tetrad $\langle L, R_S, \vec{\mathbf{D}}, \alpha \rangle$ in three-dimensional space. As shown in Figure 1.

Definition 1 (directional sensor model $\langle L, R_S, \vec{\mathbf{D}}, \alpha \rangle$). L is location of nodes, that corresponds to (x, y, z) in a three-dimensional rectangular coordinate system. R_S is maximum sensory distance of nodes. $\vec{\mathbf{D}}$ is unit vector of sensing direction, denoted by (dx, dy, dz), whose direction and central axis of sensory region is collinear. α is called sensory deviation angle and $0 \le \alpha \le \pi$.

In particular, when $\alpha = \pi$, the sensory region is a sphere thus the traditional omnidirectional sensor model can be considered as a special case of directional sensor model.

Definition 2 (probability detection model). In this part, we define the sensor sensing accuracy model. Sensing accuracy of sensor S_i at point t is defined as the probability of sensor S_i to successfully detect an event happening at point t. A point here means a physical location in the covered area.

We assume that a sensor can always detect an event happening at the point with distance 0 from the sensor, and the sensing accuracy attenuates with the increase of the distance. One possible sensing accuracy model is [24]

$$P_{it} = \frac{1}{(1 + \partial d_{it})^{\beta}}, \tag{1}$$

where P_{it} is the sensing accuracy of sensor S_i at point t, d_{it} is the distance between sensor S_i and point t, and constants ∂ and β are device-dependent parameters reflecting the physical features of a sensor. Generally, β ranges from 1 to 4. And ∂ is used as an adjustment parameter.

Sensor node density is usually higher. Assume that N sensor nodes are randomly distributed in a three-dimensional monitoring area. Therefore, events in the monitoring area are

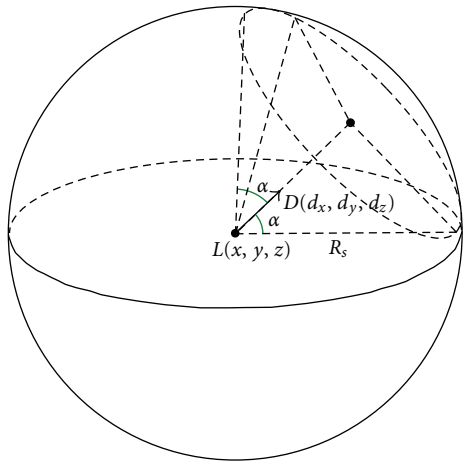


Figure 1: Directional sensor model.

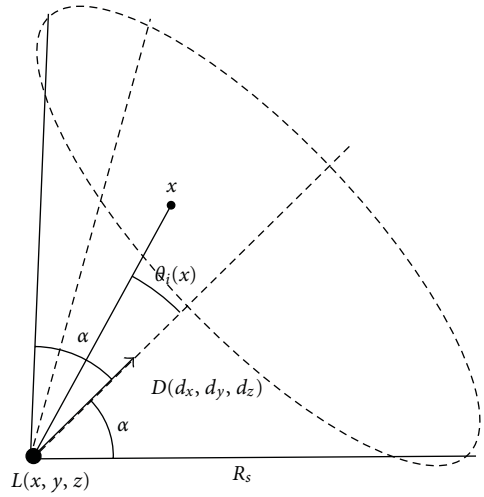


Figure 2: Sensory domain of node.

detected by multiple sensor nodes simultaneously. The sense probability is expressed as follow:

$$P_t = 1 - \prod_{i=1}^{N}(1 - P_{it}). \tag{2}$$

Substitute into Formula (1)

$$P_t = 1 - \prod_{i=1}^{N}\left(1 - \frac{1}{(1 + \partial d_{it})^{\beta}}\right). \tag{3}$$

According to (3), $P_t \geq P_{it}$, because multiple sensor nodes may sense the same events simultaneously.

Definition 3 (maximum sensory distance R_S). The maximum sensory distance of nodes is defined as

$$R_S = \frac{1}{\partial}\left(\lambda^{-1/\beta} - 1\right). \tag{4}$$

In condition of node s_i working independently, if $d_{it} \geq R_S$, the probability of node s_i being detected is $P_{it} = 1/(1 + \partial d_{it})^{\beta} \leq \lambda$. In this case, the effect on system detection probability from node s_i to target point can be ignored. λ is the minimum probability of target found, which is determined by the actual application environment, hardware and software conditions and quality of service required, and other factors. λ is usually specified by user.

Definition 4 (sensory domain of node s_i). For ease of later calculation, we translate the directional sensor model in Figure 1 into that in Figure 2 and give the following definition.

The sensory region is part of a sphere that centred on $\vec{\mathbf{D}}$ with radius R_S and maximum rotation angle which is denoted by SD_i. In other words, sensory region is constituted by all spots that satisfy Formula (1) in which x represents spots in space, θ_i represents the intersection angle with sensing direction vector $\vec{\mathbf{D}}$

$$SD_i = \{x \mid \|x - s_i\| \leq R_S, |\theta_i(x)| \leq \alpha\}. \tag{5}$$

Definition 5 (Set of neighbor nodes ψ_i). In sensor network, two nodes are called neighbors when Euclidean distance is less than twice the maximum sensory distance R_S. The set of neighbor nodes of node s_i is ψ_i

$$\psi_i = \left\{s_j \mid D\left(s_i, s_j\right) < 2R_s,\ i \neq j\right\}. \tag{6}$$

Definition 6 (γ-probability coverage). In a set of active nodes located at (x_i, y_i, z_i), $i = 1, 2, \ldots, N$, system detection probability of target point located at (x_t, y_t, z_t) is P_t. If $P_t \geqslant \gamma$, the target point located at (x_t, y_t, z_t) is satisfied with γ probability coverage. If all points in a region are satisfied with γ probability coverage, then this region is called complete γ probability coverage.

Definition 7 (degree of coverage). A given area S is equally divided into M small areas which can be assumed as points as M is large enough. If there are Q points in the M points that accord with γ-probability coverage, coverage rate of area S is defined as

$$\mathrm{DoC}(S) = \frac{Q}{M}. \tag{7}$$

Definition 8 (coverage impact factor). Coverage impact factor μ characterizes the impact on network coverage when the sensing direction angle has been changed. We define it as follows:

$$\mu = \begin{cases} 1, & \mathrm{DoC}_i'(S) > \mathrm{DoC}_i(S), \\ 0, & \mathrm{DoC}_i'(S) \leq \mathrm{DoC}_i(S). \end{cases} \tag{8}$$

$\mathrm{DoC}_i'(S)$ represents the network coverage after node s_i has been changed.

3. The Improved Virtual Potential Field-Based Algorithm in Three-Dimensional Directional Sensor Networks

Taking the deployment cost of sensor network into consideration, it would be unpractical that all nodes are capable of

moving. Moreover, the movement of sensor nodes usually causes the invalidation of part of sensor nodes and in turn changes the topology of the whole sensor network. All these factors raise the maintenance cost of network. Therefore, we assume that all nodes remain at the same location as initial and coverage can be enhanced by changing sensing direction of nodes. We introduce the concept of centroid. c_i represents the centroid that is relative to node s_i and locates in a spot in central axis of sensory region with a distance of $2R\sin\alpha/3\alpha$ apart from the node. Now the issue is translated into the virtual force issue between centroids. We assume that there is virtual repulsion $\mathbf{F}_{\text{rep}}$ between centroids. Under the action of repulsion, two nodes rotate in opposite direction to avoid the formation of sensory overlap region. At the same time of reducing the redundant coverage, a sufficient and efficient coverage of the monitoring area is achieved. Under the action of virtual potential field, every node gets the repulsion from one or more adjacent nodes.

3.1. Judgement on Overlap Situation of Sensory Area. In traditional algorithms, which use the virtual potential field to enhance coverage, the criterion for the generation of repulsion between two points is that the distance between nodes is less than $2R_S$ [7] which is applicable to omnidirectional sensor model in that when the distance between nodes is less than $2R_S$, there bound to be some overlap in sensory area. But with regard to directional sensor model, the above conclusion is obviously incorrect. As shown in Figure 3, two nodes are less than $2R_S$ apart from each other; however, with no sensory overlapped region for the difference of sensing direction angle. As a case by case, Figure 3 reflects a common situation in many cases which causes profitless adjustment of deployment, wastes energy of nodes, and shortens network lifetime.

Therefore, in this paper, the criterion for generation of repulsion between two nodes is defined as whether or not there is overlap sensory region between two nodes.

To three-dimensional directional sensor networks model, when judging whether there is overlapped sensory region between two nodes, we project the sensory region onto planes xoy, xoz, and yoz in a three-dimensional rectangular coordinate system. If and only if all projections on three planes has overlapped region, does the sensory region has overlapped region. So the overlap decision problem of three-dimensional sensor model has translated into that of two-dimensional model. I put forward an approach in [11] to decide if there is overlapped region in a two-dimensional directional sensor network model using cross-set test. Reference [25] demonstrates the 11 kinds of overlapped situations of sensory region with two-dimensional directional sensor model, as shown in Figure 4.

We simplify the sectorial sensor model of a two-dimensional space into a triangle by replacing the arc in sector by line, because only under the circumstances of (c) and (d) in Figure 4 can the decision outcome be different. It can be seen in Figure 4 that the area of the overlap region under the two circumstances is small compared to the whole sensory region, so it is inefficient to waste energy and adjust deployment of nodes for that purpose.

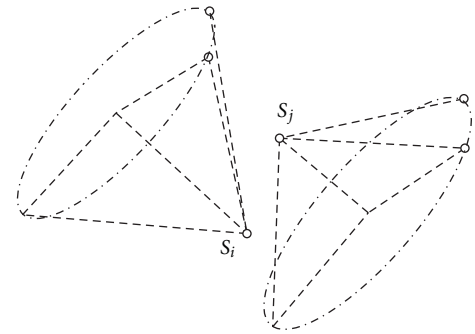


Figure 3: A case that distance is less than $2R_S$ without overlapped sensory region.

After simplifying, we can determine whether the two triangles have overlap region according to whether they have intersecting sides by mean of cross-set test. As shown in Figure 5, a triangle with overlap region must have intersecting sides. For example, in Figure 5, S_1a_1 and S_2a_2 intersect, the following must be satisfied:

$$[(a_2 - S_1) \times (a_1 - S_1)] * [(a_1 - S_1) \times (S_2 - S_1)] \geq 0,$$
$$[(S_1 - a_2) \times (S_2 - a_2)] * [(S_2 - a_2) \times (a_1 - a_2)] \geq 0. \tag{9}$$

3.2. Analysis on Regulation of Node Rotation

3.2.1. Analysis on Regulation of Centroids. Nodes s_i and s_j are neighbors, the centroid c_i at location $\mathbf{X}_i$ is under action of the repulsion of c_j at location $\mathbf{X}_j$ which is defined as follows:

$$\mathbf{F}_{\text{rep}}(i,j) = \begin{cases} \dfrac{k_{\text{rep}}}{D\left(c_i, c_j\right)^2}\left(\dfrac{\mathbf{x}_i - \mathbf{x}_j}{D\left(c_i, c_j\right)}\right), & SD_i \cap SD_j \neq \varnothing, \\ 0, & SD_i \cap SD_j = \varnothing. \end{cases} \tag{10}$$

Using the method of Section 3.1, only when there is overlap region between two nodes, repulsion between centroids of two nodes exists. $D(c_i, c_j)$ represents the Euclidean distance between centroid c_i and c_j. k_{rep} represents the repulsion coefficient which is a positive constant. SD_i/SD_j represent sensory domain of node s_i/s_j. The magnitude of repulsion of centroid is inversely proportional to the Euclidean distance between them and the direction of repulsion that the centroid c_i taking action is determined by the location of c_i and c_j. The resultant force $\mathbf{F}_{\text{rep}}(i)$ of repulsion at centroid c_i is

$$\mathbf{F}_{\text{rep}}(i) = \sum_{n_j \in \psi_i} \mathbf{F}_{\text{rep}}(i,j), \tag{11}$$

ψ_i represents the set of all the neighbors of node s_i.

3.2.2. Analysis on Rotation Angle. The resultant repulse force actions on centroid c_i and node single rotation angle θ jointly decide the later target location $\mathbf{X}_i'$ of centroid c_i. Thereby, the

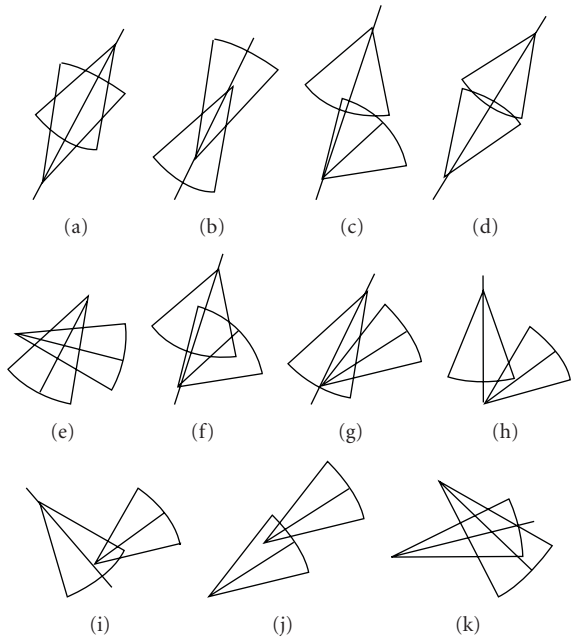


Figure 4: The situations of overlap of sensory region in directional sensor model.

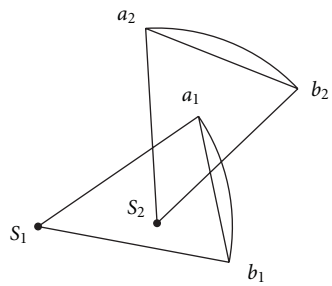


Figure 5: Two sensor models with overlap sensory region.

later target location of centroid c_i can be described as rotating a certain angle θ along the direction of resultant force $\mathbf{F}_{\text{rep}}$. Use Formula (5) to calculate the coverage impact factor. If $\mu = 1$, which means the movement is beneficial to coverage optimization, proceed the node rotation. Otherwise, the current sensing direction remain unchanged. Repeatedly, optimal solution approaches by fine adjustment. Meanwhile, we set the force threshold ε. When $\|\mathbf{F_T}\| \leq \varepsilon$, centroid will oscillate repeatedly around a certain point which can be regarded as a stable state of centroid and no more actions is required. When all centroids come to the stable state within the network, the whole sensor network is considered as having reached the stable state.

4. Algorithm Description

According to the virtual force of the centroid of each node, a node deployment adjustment algorithm is proposed as follows in this paper. The algorithm is a distributed algorithm and executes simultaneously in each node. We equally divided the sensing area S into M small areas. M is large enough so that the small areas could be assumed as points. When area degree of coverage is calculated according to formula (7), the target point should be satisfied with γ-probability coverage. Take node s_i for example, Formalized description of the algorithm is as in Algorithm 1.

5. Algorithm Simulations and Performance Analysis

We developed the simulation software Sencov3.0 that applicable to the research of sensor network deployment and coverage with VC++6.0, using which we verify the validity of IPA3D algorithm through extensive simulation experiments. Values of specific parameters are shown in Table 1. We can see from formula (4) that R_S is determined by ∂, β, λ. In all simulation experiments of this paper, we set $\partial = 1$ and $\beta = 1$, so the value of R_S is only determined by λ.

5.1. Case Study. 250 nodes are randomly deployed in a $100 * 100 * 100\,\text{m}^3$ region for monitoring the environment. The maximum sensory distance $R_S = 20$ m and sensory deviation angle $\alpha = \pi/4$. As shown in Figure 6, we record the quality of coverage in directional sensor networks as IPA3D algorithm is running in different time step. The asterisks represent sensor nodes. Due to that the simulation environment is a 3D cube; it is difficult to present the variation of global coverage intuitively from the 3D figure. Therefore, I chose to give the coverage variation of bottom side in Figure 6. During the initialization, only a few nodes cover the bottom of cube, which is shown in Figure 6(a), where the asterisks represent the nodes. Along with the running of IPA3D algorithm, some nodes that did not cover the bottom of the cube at first change their sensing directions, which results in that their sensing region now covers the bottom of the cube. At this moment, new asterisks representing those nodes appear. In a word, no extra nodes are added into that region, neither have they moved. We can see from Figure 6(a), the initial degree of coverage is 45.3%, overlap region and fade zone are significant in network for the randomness of deployment. By means of the optimization algorithm, the direction of nodes adjusts continuously as time goes on, which in consequence improves the network coverage extent, as shown in Figures 6(b) and 6(c). When the time step comes up to 60 times, the degree of coverage reaches 94.29% which increases 49 percentage points to the initial coverage, as shown in Figure 6(d).

5.2. Algorithm Convergence Analysis. We carry out a group of experiments with five kinds of network node scale so as to analyze IPA3D algorithm convergence. According to every network node scale, we randomly produce 20 topological structures, respectively, and calculate the algorithm convergence times and average. Experimental data are shown in Table 2 with parameters $R_S = 20$ m, $\alpha = \pi/4$.

Based on the above analysis we can reach a conclusion that the convergence of IPA3D algorithm, that is, the adjustment number of times, does not change conspicuously

//initialization
(1) $t \leftarrow 0$;
(2) set the maximum cycle index $t_{\max}$;
(3) determine the initial position of corresponding centroid c_i
of s_i and neighbor nodes set ψ_i.
(4) while $(t < t_{\max})$ do
(5) renew the current coordinate position of centroid c_i;
(6) determine the repulsion force $\mathbf{F}_{\text{rep}}(i)$ acting on centroid c_i from other nodes according to formula (10) and (11)
(7) if $(\|\mathbf{F}_{\text{rep}}(i)\| \geq \varepsilon)$ then
(8) determine the later target location $\mathbf{X}_i'$ of centroid c_i decided by the resultant repulse force actions on centroid c_i and node single rotation angle θ
(9) calculate the coverage impact factor μ of angle rotation this time according to Formula (8)
(10) if $\mu = 1$, proceed the angle rotation, else remain the sensing direction unchanged
(11) else break;
(12) $t \leftarrow t + 1$
(13) end.

Algorithm 1

Table 1: Experiment parameters.

Parameter name	Parameter values
Target region	$100 * 100 * 100\,\text{m}^3$
Distribution mode	Uniform distribution
Number of nodes N	100–300
Device-dependent parameter ∂, β	1, 1
Minimum probability of target found λ	0.32–0.91
γ-probability coverage	0.90
Sensory angle of deviation α	$\pi/6$, $\pi/5$, $\pi/4$, $\pi/3$, $\pi/2$, π
Repulsion coefficient k_{rep}	100

Table 2: Convergence analysis on experimental data.

Number of nodes N	Initial degree of coverage %	Ultimate degree of coverage %	Cycle index t
100	18.30	42.15	74.5
150	27.42	61.53	77.7
200	38.65	81.77	75.6
250	46.32	93.59	78.9
300	52.51	95.14	72.3

along with sensor network node scale. The value ranges from 70 to 80; thus IPA3D algorithm has a nice convergence.

5.3. Comparative Analysis of Algorithms. In this section, a series of simulation experiments are conducted to illustrate the effect on the performance of IPA3D algorithm from the three key parameters. They are node scale N, maximum sensory distance R_S, and sensory deviation angle α. Reference [9] takes traditional basis as the criterion for the generation of repulsion. We compare it to the coverage enhancement algorithm proposed in this paper and analyze their performances.

It can be seen from the curve in Figure 7 that when R_S and α are fixed, smaller value leads to less initial degree of coverage. With the increasing of node scale N, the value of Δp shows an upward trend. Δp means the difference between the final degree of coverage and the initial state. When $N = 250$, the degree of coverage increases 49 percentage points and afterwards value of Δp decreases to some extent. The reason is that when nodes number reaches a certain scale, optimized network degree of coverage has approached extreme and increase of node number can no longer enhance network coverage conspicuously. Meanwhile, the increase of nodes leads to a higher initial degree of coverage and greatly decreases the probability that several communicational adjacent nodes form coverage fade zone which undoubtedly weakens the performance of IPA3D algorithm.

We can see from the curve in Figures 8 and 9 that the effect on this algorithm from maximum sensory distance of nodes R_S and sensory deviation angle α is in accordance with the node scale. When the node scale is fixed, the smaller maximum sensory distance of nodes R_S and sensory deviation angle α are, the less possible that adjacent nodes are form an overlap region, and the less improvement is done to the network coverage performance. As the increase of maximum sensory distance of nodes, Δp increases constantly too. The network degree of coverage reaches the climax when R_S and α are at a particular value. However, with the increase of the value of R_S and α, the probability of creating coverage fade zone becomes smaller which leads to less significant effect on network degree of coverage enhancement.

As can be seen from Figures 7, 8, and 9 that compared to the PECEA algorithm in [9], under the same parameter value, the proposed IPA3D algorithm increases the coverage quality most significantly after the optimization of the initial deployment, which illustrates the superiority of this algorithm.

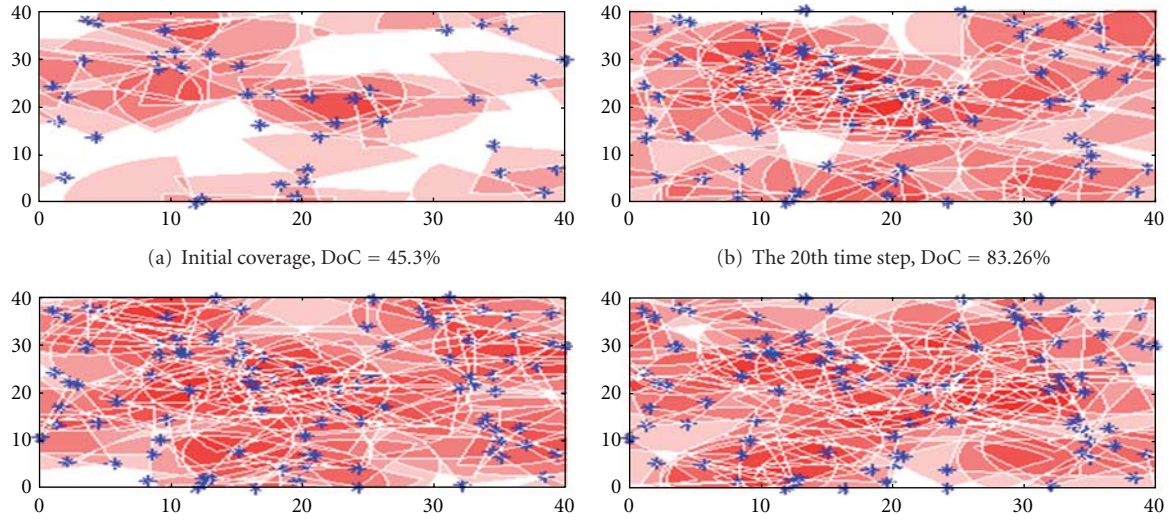


Figure 6: Coverage optimization under IPA3D algorithm.

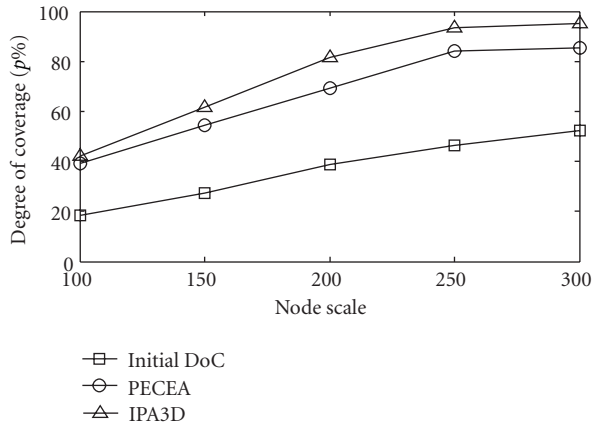


Figure 7: Effect on node scale N on condition of $R_S = 20$ m, $\alpha = \pi/4$.

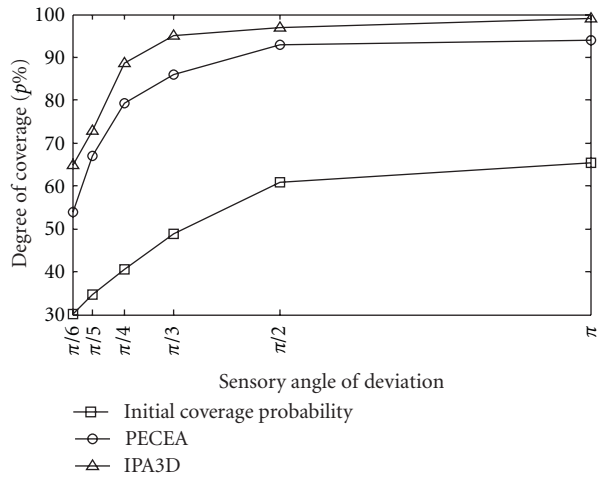


Figure 9: Effect on sensory angle of deviation α on condition of $N = 200$, $R_S = 20$ m.

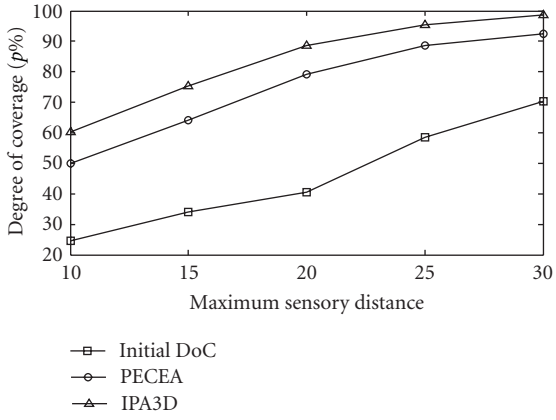


Figure 8: Effect on maximum sensory distance of nodes R_S on condition of $N = 200$, $\alpha = \pi/4$.

5.4. Analysis on Coverage Impact Factor. In this paper, we introduced coverage impact factor and use it to judge the impact on degree of coverage from the change of sensing direction angle and thereby to decide whether the node sensing direction angle should rotate. We illustrate the impact through a series of simulation experiments with $R_S = 20$ m, $\alpha = \pi/4$.

Figure 10 shows the impact from coverage impact factor on algorithm in different node scales. According to three different network node scales, 10 topologic structures are randomly generated. Record the changes of network degree of coverage when algorithm is running and average. From the curves in Figure 10 we can see, in the first 30 loops of algorithm that the degree of coverage increases sharply and exponentially. So adding coverage impact factor or not has little impact on algorithm performance. However,

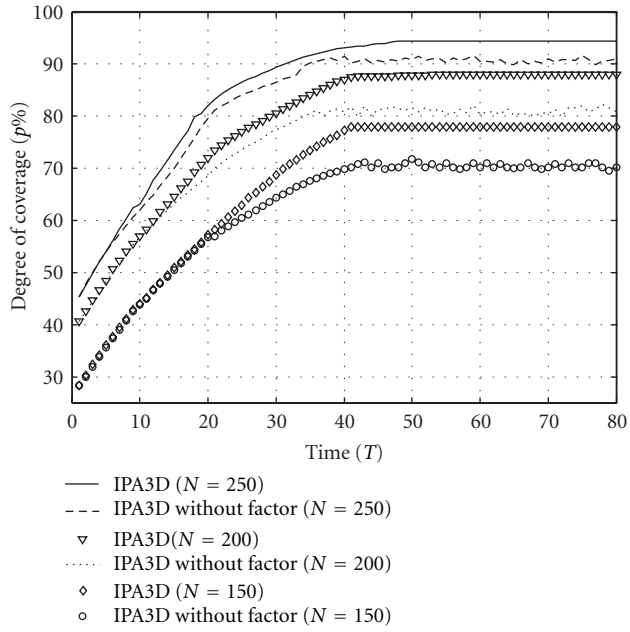


Figure 10: Impact on algorithm performance from coverage impact factor in different node scale.

with the increase of algorithm execution time, algorithm degree of coverage without coverage impact factor presents a fluctuating state and overall degree of coverage stops increasing while algorithm degree of coverage with coverage impact factor shows a gentle rising trend and finally reaches stable state with degree of coverage remaining unchanged. The reason is that, without coverage impact factor, the force on a single node is according to neighbor nodes set ψ_i and the change of sensing direction angle of a single node cannot take the impact on overall network coverage into account and may reduce network degree of coverage. And algorithm with coverage impact factor will evaluate the impact on overall coverage every time before sensing direction is changed to get rid of profitless rotation. Though it may increase algorithm complexity and cost node energy, it is greatly less than the energy cost of profitless rotation which means movements that cannot give beneficial effect to network coverage. As shown in Figure 10, to add coverage impact factor efficiently improved the performance of coverage enhancement algorithm and overcame the fault of the unstable state in later stage.

6. Conclusions

This paper proposed a probability-based three-dimensional directional sensory model, and based on which a novel criterion for judgment is proposed in view of the irrationality that traditional virtual potential field algorithms brought about on the criterion for the generation of virtual force. Cross-set test was used to determine whether the sensory region has any overlap and coverage impact factor is introduced to reduce profitless rotation from coverage optimization, thereby the energy cost of nodes was restrained and the performance of the algorithm was improved. In simulation experiment, first we verified the convergence of the algorithm and then the effect on the algorithm from the key parameters and the validity of IPA3D were demonstrated by the comparison between IPA3D and PECEA under the effect of key parameters. Also, we confirmed and analyzed the impact on IPA3D algorithm from coverage impact factor through simulation experiment. The proposed algorithm effectively improved the coverage performance of traditional virtual potential field algorithms but the energy consumption caused by the change of sensing direction angle were not taken into account which is for further study.

Acknowledgments

The subject is sponsored by the National Natural Science Foundation of China (Nos. 60973139, 61003039, 61170065, and 61171053), Scientific and Technological Support Project (Industry) of Jiangsu Province (Nos. BE2010197, BE2010198), Natural Science Key Fund for Colleges and Universities in Jiangsu Province (11KJA520001), the Natural Science Foundation for Higher Education Institutions of Jiangsu Province (10KJB520013, 10KJB520014), Academical Scientific Research Industrialization Promoting Project (JH2010-14), Fund of Jiangsu Provincial Key Laboratory for Computer Information Processing Technology (KJS1022), Postdoctoral Foundation (1101011B), Science and Technology Innovation Fund for Higher Education Institutions of Jiangsu Province (CXZZ11-0409), the Six Kinds of Top Talent of Jiangsu Province (2008118), Doctoral Fund of Ministry of Education of China (20103223120007, 20113223110002), and the Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (yx002001).

References

[1] S. Meguerdichian, F. Koushanfar, M. Potkonjak, and M. B. Srivastava, "Coverage problems in wireless ad-hoc sensor networks," in *Proceedings of the 20th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '01)*, pp. 1380–1387, IEEE Press, New York, NY, USA, April 2001.

[2] D. Tao, Y. Sun, and H. J. Chen, "Worst-case coverage detection and repair algorithm for video sensor networks," *Acta Electronica Sinica*, vol. 37, no. 10, pp. 2284–2290, 2009.

[3] I. F. Akyildiz, T. Melodia, and K. R. Chowdhury, "A survey on wireless multimedia sensor networks," *Computer Networks*, vol. 51, no. 4, pp. 921–960, 2007.

[4] A. Howard, M. J. Mataric, and G. S. Sukhatme, "Mobile sensor network deployment using potential field: a distributed scalable solution to the area coverage problem," in *Proceedings of the 6th International Symposium on Distributed Autonomous Robotics Systems (DARS '02)*, pp. 299–308, Fukuoka, Japan, June 2002.

[5] S. Poduri and G. S. Sukhatme, "Constrained coverage for mobile sensor networks," in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 165–171, IEEE Press, New York, NY, USA, May 2004.

[6] Y. Zou and K. Chakrabarty, "Sensor deployment and target localization in distributed sensor networks," *ACM Transactions on Embedded Computing Systems*, vol. 3, no. 1, pp. 61–91, 2004.

[7] J. Ai and A. A. Abouzeid, "Coverage by directional sensors in randomly deployed wireless sensor networks," *Journal of Combinatorial Optimization*, vol. 11, no. 1, pp. 21–41, 2006.

[8] S. J. Li, C. F. Xu, Z. H. Wu, and Y. H. Pan, "Optimal deployment and protection strategy in sensor network for target tracking," *Acta Electronica Sinica*, vol. 34, no. 1, pp. 71–76, 2006.

[9] D. Tao, H. D. Ma, and L. Liu, "A virtual potential field based coverage-enhancing algorithm for directional sensor networks," *Journal of Software*, vol. 18, no. 5, pp. 1151–1162, 2007.

[10] Y. L. Cai, W. Lou, M. L. Li, and X. Y. Li, "Target-oriented scheduling in directional sensor networks," in *Proceedings of the 26th IEEE International Conference on Computer Communications (INFOCOM '07)*, pp. 1550–1558, Anchorage, Alaska, USA, May 2007.

[11] J. J. Huang, Y. Shi, L. J. Sun, R. C. Wang, and H. P. Huang, "An improved algorithm of virtual potential field in directional sensor networks," *Key Engineering Materials Journal*, vol. 467–469, pp. 1562–1568, 2011.

[12] H. D. Ma and Y. H. Liu, "Correlation based video processing in video sensor networks," in *Proceedings of the IEEE International Conference on Wireless Networks, Communications and Mobile Computing (WirelessCom '05)*, pp. 987–992, Beijing, China, June 2005.

[13] H. D. Ma and D. Tao, "Multimedia sensor network and its research progresses," *Journal of Software*, vol. 17, no. 9, pp. 2013–2028, 2006.

[14] M. C. Zhao, J. Y. Lei, M. Y. Wu, Y. Liu, and W. Shu, "Surface coverage in wireless sensor networks," in *Proceedings of the 28th IEEE Conference on Computer Communications (INFOCOM '09)*, pp. 109–117, Rio de Janeiro, Brazil, April 2009.

[15] J. Adriaens, S. Megerian, and M. Potkonjak, "Optimal worst-case coverage of directional field-of-view sensor networks," in *Proceedings of the 3rd Annual IEEE Communications Society on Sensor and Ad hoc Communications and Networks (Secon '06)*, pp. 336–345, September 2006.

[16] M. Cardei, M. T. Thai, Y. Li, and W. Wu, "Energy-efficient target coverage in wireless sensor networks," in *Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '05)*, pp. 1976–1984, Boca Raton, Fla, USA, March 2005.

[17] Z. Xiao, M. Huang, J. H. Shi, J. Yang, and J. Peng, "Full connectivity and probabilistic coverage in random deployed 3D WSNs," in *Procedings of the International Conference on Wireless Communications and Signal Processing, (WCSP '09)*, vol. 31, no. 12, pp. 1–4, Kunming, China, November 2009.

[18] J. J. Huang, L. J. Sun, R. C. Wang, and H. P. Huang, "Virtual potential field and covering factor based coverage-enhancing algorithm for three-dimensional wireless sensor networks," *Journal on Communications*, vol. 31, no. 9, pp. 16–21, 2010.

[19] H. D. Ma, X. Zhang, and A. L. Ming, "A coverage-enhancing method for 3D directional sensor networks," in *Proceedings of the 28th IEEE Conference on Computer Communications (INFOCOM '09)*, pp. 2791–2795, Rio de Janeiro, Brazil, April 2009.

[20] D. X. Zhang, M. Xu, Y. W. Chen, and S. Wang, "Probabilistic coverage configuration for wireless sensor networks," in *Proceedings of the International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM '06)*, pp. 1–4, Changsha, China, September 2006.

[21] J. P. Sheu and H. F. Lin, "Probabilistic coverage preserving protocol with energy efficiency in wireless sensor networks," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '07)*, pp. 2631–2636, Kowloon, China, March 2007.

[22] X. Bai, C. Zhang, D. Xuan, and W. Jia, "Full-coverage and k-connectivity (k = 14, 6) three dimensional networks," in *Proceedings of the IEEE Annual Conference on Computer Communications (INFOCOM '09)*, pp. 145–154, Rio de Janeiro, Brazil, April 2009.

[23] S. M. N. Alam and Z. J. Haas, "Coverage and connectivity in three-dimensional networks," in *Proceedings of the 12th Annual International Conference on Mobile Computing and Networking (MOBICOM '06)*, pp. 346–357, Los Angeles, Calif, USA, September 2006.

[24] J. Lu and T. Suda, "Coverage-aware self-scheduling in sensor networks," in *Proceedings of the Annual Workshop on Computer Communications (CCW '03)*, pp. 117–123, IEEE, California, Calif, USA, October 2003.

[25] D. Pescaru, C. Istin, D. Curiac, and A. Doboli, "Energy saving strategy for video-based wireless sensor networks under field coverage preservation," in *Proceedings of the IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR '08)*, vol. 1, pp. 289–294, Cluj-Napoca, Romania, May 2008.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 920505, 9 pages
doi:10.1155/2012/920505

Research Article

An ACOA-AFSA Fusion Routing Algorithm for Underwater Wireless Sensor Network

Huafeng Wu,[1] Xinqiang Chen,[1] Chaojian Shi,[1] Yingjie Xiao,[1] and Ming Xu[2]

[1] *Merchant Marine College, Shanghai Maritime University, Shanghai 201306, China*
[2] *College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China*

Correspondence should be addressed to Huafeng Wu, hfwu@shmtu.edu.cn

Received 10 October 2011; Revised 3 March 2012; Accepted 4 March 2012

Academic Editor: Sabah Mohammed

Due to intrinsic properties of aqueous environments, routing protocols for underwater wireless sensor network (UWSN) have to cope with many challenges such as long propagation delay, bad robustness, and high energy consumption. Basic ant colony optimization algorithm (ACOA) is an intelligent heuristic algorithm which has good robustness, distributed computing and combines with other algorithms easily. But its disadvantage is that it may converge at local solution, not global solution. Artificial fish swarm algorithm (AFSA) is one kind of intelligent algorithm that can converge at global solution set quickly but has lower precision in finding global solution. Therefore we can make use of AFSA and ACOA based on idea of complementary advantages. So ACOA-AFSA fusion routing algorithm is proposed which possesses advantages of AFSA and ACOA. As fusion algorithm has aforementioned virtues, it can reduce existing routing protocols' transmission delay, energy consumption and improve routing protocols' robustness theoretically. Finally we verify the feasibility and effectiveness of fusion algorithm through a series of simulations.

1. Introduction

Rapid evolvement of wireless communication technology, electronic technique, sensor technology, micro-electromechanical system, and other computer technologies promotes research in wireless sensor network. Underwater wireless sensor network (UWSN), one kind of special and promising wireless sensor network, gains more and more attention. With the development of ocean exploitation, UWSN is becoming one of hot topics in research fields. UWSN is the development trend of future underwater communication and detection techniques, for instance, seabed resource detection and exploration, oil spill, tsunami, underwater earthquake, and underwater environment monitoring. Therefore we can say that marine development degree determines future world development degree while UWSN determines marine development.

Electromagnetic wave and light wave are not suitable for underwater communication since their signals can be absorbed by water. As wavelength of acoustic wave is long enough and it is cost effective, acoustic communication is the only ideal medium for underwater information transmission. That is to say, acoustic communication is the best communication style for UWSN to date [1]. However the transmission condition of ocean acoustic channel is terrible, and there are so many interference factors that disturb such information transferring [2]. Besides UWSN is disposed in interesting aquatic area which means nodes of UWSN are mobile and delicate easily. And above all, each node of UWSN carries limited energy and it is impossible to replace its battery. All of those afore-mentioned characteristic are huge challenge for the realization of UWSN. As most traditional land-based wireless sensor networks routing protocol and other energy-based routing protocols are designed for stationary network, they are not applicable to UWSN whose network topology is dynamic. Therefore UWSN routing protocol is becoming a hot spot in the field. The design principle of UWSN routing protocol is the same as WSN—low energy consumption [3]. UWSN is a kind of three-dimensional dynamic, sparse networks, so it is important to take such characteristics into consideration in devising underwater routing protocol. In other words,

underwater routing protocol algorithm should have these following advantages: energy efficiency, short transmission path, and robustness.

Ant colony optimization algorithm (ACOA) is one kind of heuristic bionic algorithm based on ant colony finding its way in the population's foraging process. ACOA algorithm is parallel algorithm in essence that has the following features: robustness, universality, fast convergence, easiness of combining with other algorithm, and so forth [4, 5]. However ACOA may converge at local optimal solution for the algorithm refreshes local pheromones, so some improvements need to be taken to optimize ACOA for its application in UWSN. Compared with ACOA, AFSA converges at global solution better and is insensitive to parameters [6, 7]. So we can take advantage of AFSA to optimize ACOA, and apply the fusing algorithm based on ACOA, and AFSA to the routing protocol of UWSN, which can improve protocol's robustness and energy efficiency.

The remainder of the paper is organized as follows. In Section 2 we present related work. In Section 3, we present AFSA-ACOA fusion algorithm which includes introduction of ACOA, AFSA, the ACOA-AFSA fusion algorithm, the implementation, and its application in optimizing routing protocol of UWSN. Section 4 mainly gives details for the simulation and analysis of the simulation result. At last Section 5 concludes the paper.

2. Related Work

Lv et al. [8] provided a research review of mobile underwater wireless sensor network. It included the following contents: node's design of underwater wireless sensor network, nodes interconnection and dynamic network, nodes' location in three-dimensional environments, and acoustic channel link control. But Lv et al. did not pay much attention to UWSN's routing protocol, not to say the research about its optimization making use of some algorithms. Guo et al. [9] did the similar work as Lv et al. Meanwhile Guo introduced some research work with respect to routing protocol of UWSN. He said underwater routing protocol could be divided into three categories: initiative routing protocol, on-demand routing protocol, and routing based on geography. But Guo et al. did not present research about optimal algorithms for UWSN's routing protocol.

Yanhua [1] proposed time-sharing balance routing algorithm based on network coding for UWSN. It is one kind of improvement flooding algorithm. The routing protocol had the following characteristics: (a) children nodes need not to response to their father nodes in routing establishment process; they would broadcast their routing data packets which could be considered as acknowledgment to their father's requirement. Such activity saved time and energy effectively. What is more, the action reduced conflicts probability in sending data packets significantly. (b) Probability balance mechanism was given to regulate every communication cluster's children nodes. Thus each node's overhead in the routing tree was equivalent. (c) After establishment of routing tree, data packets received by nodes were encoded into one encoding, compressed data packet. Such encoding technique reduced node's energy cost and improved network bandwidth's utilization ratio, prolonged network's lifetime. Xiao et al. [10] reported the ongoing efforts of their research toward developing an analytic model to address the performance of contention-based protocols within the context of underwater acoustic sensor networks. They identified the challenges of modeling contention-based MAC protocols and present models for analyzing ALOHA and p-persistent ALOHA variants for a simple string topology.

Qiliany and Liping [11] took sensor nodes' limited energy, communication into account, and he proposed one routing protocol based on improved optimal ant colony algorithm. The algorithm considered node's residual energy, communication distance, and other factors fully in routing list; it gathered ants' searching activities near optimal solution successfully. To avoid accident of premature convergence, the algorithm restricted range of pheromones. But the algorithm is suitable for land-based wireless sensor network. Similarly, Xuhua et al. [12] discussed a new routing approach based on ant colony optimization algorithm to obtain the optimal path between two nodes in network. The algorithm increased formulas' sensibility to impact factors, thus the algorithm's convergence speed was raised largely. Xiangguang et al. [13] posed a new hybrid evolutionary algorithm based on artificial fish swarm algorithm and particle swarm optimization algorithm. The core is such algorithm took advantage of ASFA's global convergence to find range of satisfied solution firstly, then it employed PSO algorithm to get local optimal solution quickly. So the hybrid method had features of quick local searching speed and global convergence. For the multiple QoS constrained unicast routing problem, a new QoS routing algorithm combining modified ant colony algorithm with artificial fish swarm algorithm was proposed by Mingjia et al. [14]. It adopted hybrid ant behavior to produce diverse original paths, optimizing the choice nodes set according to multiple QoS constrained, adding AFSA to MACA's every generation, making use of AFSA's advantage of whole quick convergence, ACA's convergence speed was quickened, and AFSA's preying behavior improved the ability of MACA to avoid being premature.

Muhammad Ayaz and Abdullah [15] proposed a Hop-by-Hop Dynamic Addressing Based (H2-DAB) to provide scalable and time-efficient routing for UWSN. The H2-DAB routing does not require any dimensional location information or any extra specialized hardware compared with many other routing protocols in the same area. Domingo and Prior [16] presented GPS-free Routing Protocol for UWSN in deep water. It minimized the proactive routing message exchange and compensated the high propagation delays of the underwater medium using a continually adjusted timing advance combined with guard time values to minimize data loss and maintain communication quality. Huang et al. [17] proposed a self-healing clustering algorithm which combined the ideas of energy-efficient cluster-based routing and application-specific data aggregation. The self-healing mechanism significantly enhanced the robustness of clustered UWSNs.

The aforementioned researches are closely related with sensor network routing and its algorithms. As characteristics of subaqueous environments, routing algorithms with short propagation delay, good robustness, low energy consumption, and high network throughput can be really challenged for UWSN. To address the above-mentioned challenges, we propose one hybrid ACOA-AFSA routing algorithm for underwater wireless sensor network.

3. ACOA-AFSA Fusion Algorithm and Its Application in Optimization of UWSN's Protocol

3.1. Introduction of ACOA and AFSA. Ant will leave pheromones in their passing roads so that other ants can find the previous ants easily. What is more, ant is going to take the path which has higher pheromones, left by other ants, than other pathways nearby. So ant colony can find their end readily with the help of pheromones' positive information feedback. That's to say, although ant is nonintelligent species, activities of ant colony represent living intelligent. Advantages of ACOA are self-organization, distributed process, positive feedback, and good robustness. Thus ACOA is suitable for solution of NP-hard problems like routing protocol in UWSN. But we notice that ACOA may converge at local optimal solution instead of global optimal solution sometimes.

Fish can find specific areas that are rich in nutrients in their living waters by themselves or tagging along with other fish quickly. Thus the area that has the greatest number of fish is the most nutritious area in the water. According to the characteristics, artificial fish swarm algorithm is proposed by Xiaoli [6]. Artificial fish swarm algorithm (AFSA) can get global solution effectively by simulating activities of fish swarm's foraging, clustering, and rear ending. AFSA has simple algorithm model and strong abilities of getting out of local solution, converging at global solution. However the AFSA can hardly get high-accuracy solution sometimes. Besides in the foraging stage of AFSA, individual artificial fish may choose some new status randomly if the fish cannot get better status than current state. That means previous beneficial information is not used fully. Thus advantages and disadvantages of AFSA can be used and made up by ACOA.

3.2. Introduction of Fusion Algorithm. Basic idea of the fusion algorithm is that fusion algorithm takes artificial fish swarm algorithm as subject, and it introduces idea of ant colony optimization algorithm. So ant colony algorithm takes advantage of AFSA's speediness and global solution to fulfill quick convergence and get global solution. At the same time ACOA covers disadvantages of AFSA. As ACOA has strong ability of positive feedback for ants search for pheromones left by previous ants fully, every next state of ACOA will be better than its current status. Such superiority can cover disadvantage of AFSA effectively. With ACOA's calibration, AFSA can modify its routing path more slightly and accurately.

3.3. Implementation of Fusion Algorithm See [18]. AFSA can get optimal solution domain quickly with lower precision. Meanwhile an important feature of ACOA is that the algorithm makes use of pheromones' positive feedback to choose optimal solution. Thus based on the idea of offsetting each other's weakness, ACOA-AFSA fusion algorithm is proposed which can converge at optimal solution effectively and quickly. To summarize the realization of the proposed fusion routing algorithm, fish swarm tries to find routing path through data delivering from source node to destination node. And each node in the routing path then compares its energy, path length with nearby nodes with the assistance of ACOA algorithm. Algorithm's basic steps are as follows while its flow diagram is shown in Figure 1.

(a) Parameters' initialization: both parameters of ant colony algorithm and AFSA are initialized. They are shown as follows. Parameters in ant colony algorithm: pheromone evaporation factor ρ, pheromone level Q, number of ants K, maximum iteration number for ant colony algorithm Max, pheromone heuristic factor α. Parameters in AFSA: fish's step length S, visual radius and bulletin board's content, population quantity, crowding factor δ, and so on.

(b) i, identifier number of fish, increments one.

(c) Implement foraging activity. Fish swarm implements activity of foraging to test whether next status is better than current status. If so, the fish swarm implements foraging activity. If not AFSA will perform next activity: clustering activity. The logic formula of environment consistence is shown

$$x(i)_{\text{consistence}} = \left(\frac{\sin(x_{i_1})}{x_{i_1}}\right) * \left(\frac{\sin(x_{i_2})}{x_{i_2}}\right), \tag{1}$$

where x_i is current state of fish i and x_{i_1}, x_{i_2} are the coordinates of x_i, respectively, $x(i)_{\text{consistence}}$ is consistence of x_i. Fish i implements activity of foraging by comparing the consistence with its neighbor. In its visual range, fish i chooses some status randomly and the chosen status is compared with current status to judge which one is better. If the chosen one is better, then fish i will move one step toward it. Else fish will try to find some better status in the extent of maximum number of iterations. The equations are shown as follows:

$$x_v = x_i + \text{rand}() \times \text{visual} \tag{2}$$

$$x_{\text{inext}} = x_i + \text{rand}() \times \text{step} \times \frac{x_v - x_i}{\|x_v - x_i\|}, \tag{3}$$

where x_v is the chosen state that needs to be compared and x_{inext} is the next state that fish i will reach. Here (2) is valid if fish i finds better status in the maximum iterations, and (3) is executed for fish's next step toward x_v. After that every status' consistence will be increased with idea of ACOA. That is to say, we introduce modifying factor, which is positive feedback mechanism from AOCA, to strengthen the chosen status (next node). The updated consistence is shown as expression (4). Then bulletin board's content is refreshed by latest status. We employ two pieces of bulletin boards to record fish's choice. One is for recording the best

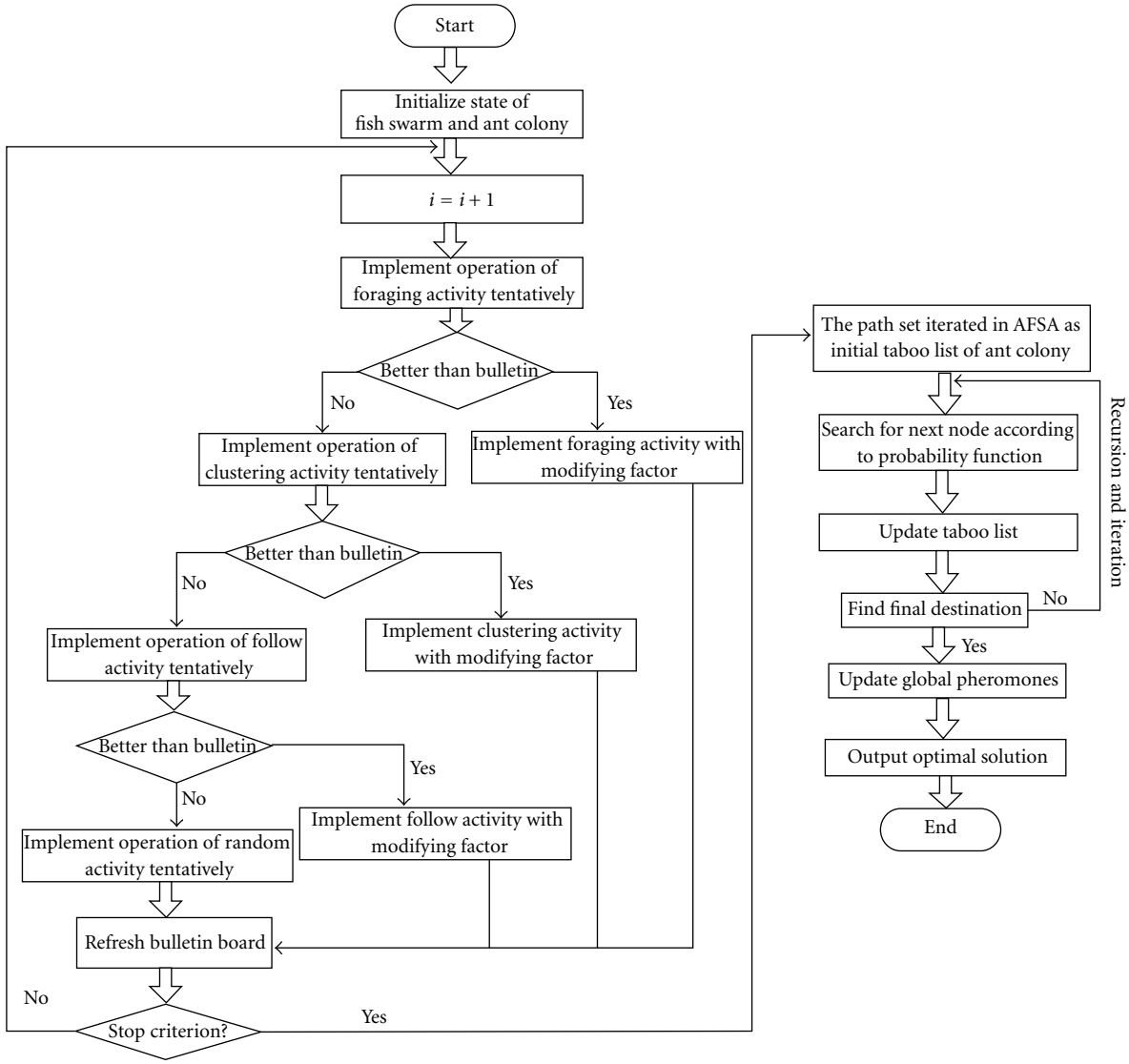


Figure 1: Fusion algorithm's flow diagram.

optimal choice and the other one is for suboptimal choice. The content of the bulletin boards will be refreshed after every activity. We name the boards as optimal board and sub-optimal board, respectively. If there is only one link for fish to choose or the distance of alternative node is bigger than optimal path, then the content of optimal board is the same as sub-optimal board. The rules are applicable to the clustering activity, following activity and random activity:

$$x(\text{inext})_{\text{consistence}} = x(\text{inext})_{\text{consistence}} + \text{rand}() \times \frac{x(\text{inext})_{\text{consistence}}}{x(i)_{\text{consistence}}}. \quad (4)$$

(d) Implement clustering activity. fish swarm implements activity of clustering to test whether next status is better than current status. If so, the fish swarm implements foraging activity. If not AFSA will perform next activity: following activity. The activity of clustering is similar with foraging activity. First fish i percepts the number of fish in its visual range and the fish's visual range forms set of visual_fish s_i. Center location of fish is calculated in (6):

$$s_i = \left\{ x_i \mid \left|\left| x_j - x_i \right|\right| \le \text{visual}, j = 1, 2, 3, \ldots, n \right\}, \quad (5)$$

$$x_{\text{center}} = \frac{\sum_{j=1}^{n} x_j}{n}. \quad (6)$$

Fish i will implement clustering activity if those two conditions are met: (1) $x(\text{center})_{\text{consistence}} < x(i)_{\text{consistence}}$, (2) $\|x(j) - x(i)\| < \delta$, $j \neq i$, $j = 1, 2, \ldots, n$. The two premises demonstrate that $x(\text{center})$ is better than $x(i)$. Logically the next status of fish i is shown in expression (7). Like foraging activity, consistence of next status should be strengthened with (4) and then bulletin board's content is refreshed by latest status:

$$x_{\text{inext}} = x_i + \text{rand}() \times \text{step} \times \frac{x_{\text{center}} - x_i}{\|x_{\text{center}} - x_i\|}. \quad (7)$$

(e) Implement following activity: fish swarm implements activity of following to test whether next status is better than current status. If so, the fish swarm implements foraging activity. If not AFSA will perform next activity: random activity. The following activity is an activity that fish tries to get closer to fish state that has minimum consistence. Certainly following activity will be implemented if the following conditions are met:

(1) $x(\text{minimum})_{\text{consistence}} < x(i)_{\text{consistence}}$,

(2) $\|x(\text{center}) - x(i)\| < \delta$.

Similarly the two conditions indict that $x(\text{minimum})$ is better than $x(i)$, so fish i will move one step toward x (minimum). The next state of fish i is shown

$$x_{\text{inext}} = x_i + \text{rand}() \times \text{step} \times \frac{x_{\text{minimum}} - x_i}{\|x_{\text{minimum}} - x_i\|}. \quad (8)$$

So the same work is done to improve environment consistence like foraging activity and clustering activity. At last AFSA updates content of bulletin board.

(f) Random activity: if none of foraging activity, clustering activity, and following activity is implemented, then random activity is executed. The calculation formula for next step of fish is shown in expression (9). As shown in flow diagram, next step is to judge whether fish swarm gets optimal solution or AFSA reaches maximum iterations. If not, algorithm jumps to step b and implement circle successively:

$$x_{\text{inext}} = x_i + \text{rand}() \times \text{step}. \quad (9)$$

(g) ACOA optimizes the path set from AFSA. Although AFSA has strong ability of finding optimal solution set, its disadvantage is that it has lower precision of obtaining optimal solution in optimal solution set. Because of the aforementioned reasons, we use ACOA to acquire optimal solution from the solution set that outputted from AFSA. Firstly taboo list in ACOA is initialized by solution set of AFSA and ant colony is the same as fish swarm.

(h) Searching for next node and transferring path: we use taboo list $\text{Tab}u_k$ to record the city that ant k walks around. And the transferring path is determined by route pheromone heuristic index α. P_{ij}^k represents probability that ant k moves to city j from city i. Its expression is shown in the following term:

$$P_{ij}^k = \begin{cases} \dfrac{\left[\tau_{ij}(t)\right]^{\alpha} \cdot [\eta_{ik}(t)]^{\beta}}{\sum_{s \subset \text{allowed}_k} [\tau_{is}(t)]^{\alpha} \cdot [\eta_{is}(t)]^{\beta}}, & j \in \text{allowed}_k, \\ 0, & \text{else.} \end{cases} \quad (10)$$

Parameter allowed_k in expression (10) indicates potential cities that ants can choose for the next step. $\eta_{ij}(t)$ indicates heuristic function and its expression is as follows:

$$\eta_{ij}(t) = \frac{1}{d_{ij}}, \quad (11)$$

where d_{ij} is the distance between city i and j.

```
{ X = afinit (afNum,varNum,Xl,Xh,lchrom);
%initialization of AFSA
for i= 1 : afNum
    temp=[X(i,1),X(i,2)];
    Tempconsistence(i)=functionconsistence(temp,
    sim_model);
end
foodconsistence=tempconsistence';
% calculate food density
[a,b]=max(foodconsistence);
Xmax=[X(b,1),X(b,2)];
Ymax=a;
%initialize bulletin board
k=0;
visualweak=0;
while(k<iterativeTime)
for i=1:afNum
Xi=[X(i,1),X(i,2)];
%current status of fish i
nf1=0;
Xc=0; }
```

Algorithm 1

(i) Path pheromone update. Pheromones level in each path that ants have passed by is refreshed in specific algorithm to induce more ants select the path. The update policy as follows:

$$\Delta\tau_{ij}^k(t) = \begin{cases} \dfrac{QN}{L_K^2}, & \text{ant } k \text{ pass city } i \text{ and } j \\ 0, & \text{else.} \end{cases} \quad (12)$$

In formula (12), Q represents pheromones' level; it affects convergence speed to some extent. L_K^2 represents the square of path length that ant k walks by in the circulation. N means total number of nodes that ant k has to traverse from source node to destination node [19, 20].

(j) Loop condition judgment: if whole ant colony walks around the same and the only one route to destination, or number of iterations reaches the maximum iterations set by the algorithm, the ant colony algorithm stops. Otherwise algorithm will jump to step h and continue the search for next path. Pseudo-code block of fusion algorithm's about AFSA's initialization as is in Algorithm 1.

3.4. Issue Description and Solution. Usually UWSN consists of many underwater nodes for monitoring underwater environments. But characteristics of underwater environments, such as long propagation delay, multipath interference, and limited channel width, decrease UWSN's network performance. Nodes and network link may join in UWSN or be unusable at any time. Namely, protocols of UWSN may have bad robustness, higher energy consumption, longer transmission delay, lower throughput of network, and so forth.

Thus efficient routing protocol algorithm for UWSN should resolve the aforementioned deficiencies effectively.

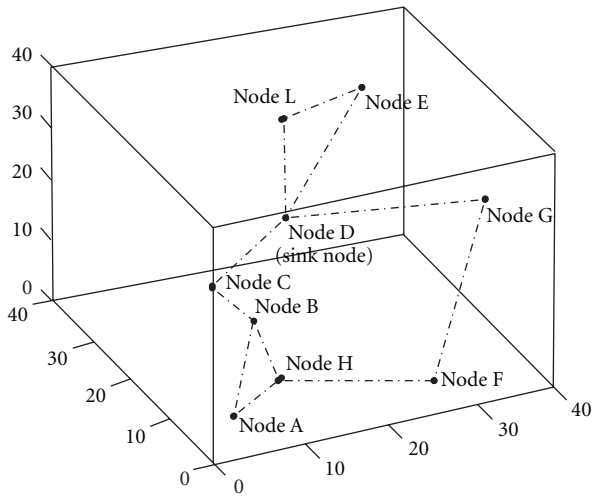


Figure 2: Topology link of UWSN.

The conspicuous characteristics of efficient routing algorithm are that the algorithm has autocatalysis and positive feedback mechanism. Fusion algorithm proposed in this paper is the very algorithm that possesses the aforementioned characteristics. As fusion algorithm is based on ant colony algorithm and artificial fish swarm algorithm, the fusion algorithm has advantages of both algorithms like positive feedback mechanism, self-adaptive mechanism, converging at global solution quickly, and good robustness.

Fusion algorithm's self-adaptive mechanism enhances robustness of UWSN's routing protocol. In this case routing protocol will not be affected if partial nodes become invalid or some network links are disabled suddenly. Thus robustness of routing protocol is strengthened. What is more, the positive feedback mechanism makes message routes from source node to destination node in shortest path. That means overhead of information routing will be decreased naturally. The more nodes deployed in acoustic environment and further distance that message need to be delivered, the more overhead will be saved. Taking all the aforementioned contents into consideration, we can employ the fusion algorithm to improve UWSN's network robustness, reduce its energy consumption and propagation delay. Because of reasons given above, we will use ACOA-AFSA fusion algorithm for finding optimal link in UWSN.

For simplicity, we give communication links in protocol table of UWSN as shown in Figure 2. Node A is the source node and node D is sink node while node E is destination node. The distance between AB, BC, CD, DE, AH, HF, FG, GE, BH, LE, LD, DG is 10, 6, 10, 12, 8, 15, 18, 17, 5, 5, 9, 5, 16, respectively. Firstly fish swarm computes food density in the environment. That is to say, each fish in the population will implement foraging activity, clustering activity, following activity, or random activity to choose next node to jump according to formulas from (1) to (9). So after the fish swarm finishes the first circle, most fish will choose node H and node B as their next node, while only little fish will choose node I as their next node. So the optimal bulletin will record H and sub-optimal board will record B as their content. Then the fish swarm continues to find next node based on their current position until they arrive at node D which is a sink node. Next we extract the content in two bulletin boards. The path in optimal-board is A-H-B-C-D while the path in sub-optimal board is A-B-C-D. Then we use ACOA to choose which path is the real shortest path. Paths in optimal board and sub-optimal board form solution set. The source node and destination node in ACOA are node A and node D, respectively. Then ants walk along the given routes from source node to destination node. As The length of path A-B-C-D is 26 and the length of path A-H-B-C-D is 29. As a result, more and more ants will choose path A-B-C-D than path A-H-B-C-D as their route to destination. The whole ant colony will choose A-B-C-D as their walking path eventually. And the process of choosing path from D to E is similarly with choosing path from A to D. So the final routing path is A-B-C-D-E and the whole length of the path is 38.

4. Simulation and Analysis

4.1. Simulation Setup. We ran series of simulations to evaluate performance of the proposed algorithm by comparing with other popular algorithms which were discussed in Section 3.2 concretely. In the series of experiment, 1000 nodes are randomly deployed in underwater environment in $1000 \times 1000 \times 500$ cubic meters which includes 80 sink nodes. The communication range for sensor nodes and sink nodes was 60 meters while data channel was set to 5 kbs. The length of each data message packet is 512 bytes. Based on the energy consumption model presented in [17, 21], the least transmission power required at the transmitter site to achieve a power level, p_0, at the receiver can be expressed as follows:

$$E_{tx}(d) = p_0 \times d^k \times 10^{d\times(\alpha(f)/10)}, \tag{13}$$

where d is the distance between transmitting node and receiving node, k is loss factor that is affected by extension of wave surface, $\alpha(f)$ is absorption coefficient of frequency, while f is the acoustic transmitting frequency. The parameters of AFSA and ACOA are as follows. (1) Parameters in AFSA. Fish's step length S is 2, bulletin board's content is 0, visual radius is 6, crowding factor δ is 10, iteration number is 2000, calculation precision is 10^{-3}, and population quantity of fish swarm is 50 [7]. (2) Parameters in ACOA. Pheromone evaporation factor ρ is 0.5, number of ants is 50, maximum iteration number for ant colony algorithm Max is 1000, pheromone heuristic factor α is 1.5, expected heuristic factor β is 5, and strength of pheromone Q is 600 [22, 23].

4.2. Simulation Scheme. The simulation experiments were realized through C++ to verify feasibility and validity of the proposed algorithm. The compared schemes during the simulations included the proposed ACOA-AFSA fusion algorithm, a representative clustering algorithm called low-energy adaptive clustering hierarchy (LEACH) [24], classic routing protocol for underwater sensor networks named vector-based forwarding (VBF) [25].

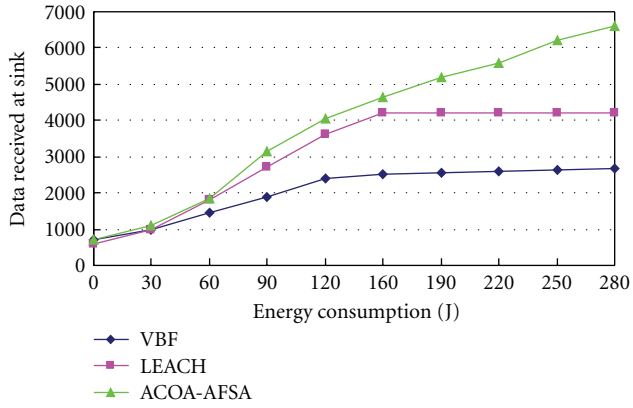


Figure 3: energy consumption versus data receive at sink node.

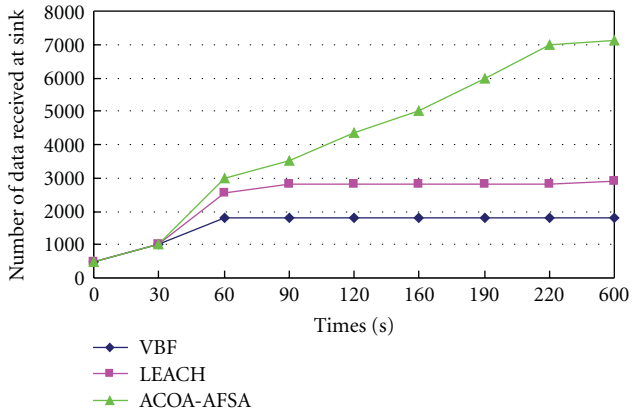


Figure 4: Total data packets received at sink node.

We compared the three algorithms in the following aspects. (1) energy consumption: we compared energy consumption in the same application and condition, such as same nodes, same velocity, same packets need to be delivered. (2) Loss ratio of data packets: we got each algorithm's packets loss probability through quantity loss of the sent packets and received packets between same source node and destination node in same application and condition, such as same nodes, same velocity, and same packets that need to be delivered. (3) Propagation time and delay: propagation time and delay were measured by transmission time of data packets from same source node and destination node in the same application and condition such as same nodes, same velocity, same packets need to be delivered.

4.3. Simulation Results and Analysis

(a) Energy Consumption Comparison. Figure 3 shows the number of data received at sink nodes over different energy consumption level. In fact both algorithms of VBF and LEACH have intrinsic drawback of limited receiving capacity. Figure 3 reveals that LEACH protocol's largest capacity is 4200 approximately while VBF is 2500 data packets. The proposed routing protocol in the paper does not have such imperfection. Mainly reasons for the aforementioned phenomenon can be summed up as follows: (a) LEACH protocol assumes the nodes chosen as clustering randomly without considering energy state, and it presumes cluster nodes are scattered in monitoring area evenly. As a matter of fact, not all of the message collected by nodes in cluster can be sent to sink nodes correctly and effectively in the real and complicated monitoring area. Sometimes part of monitoring area is full of monitoring nodes but no clustering nodes, thus the information collected by spot nodes cannot be transferred and relayed in time. As a result, energy consumption of LEACH protocol is proportional to the amount packets while data packets fall short of protocol's theoretical maximum. LEACH's energy consumption increases while its packets remain unchanged if protocol reaches its maximum limitation. (b) In basic VBF protocol, nodes that near sink nodes think they are authorized to deliver and relay data packets. Thus same data packets may be transferred several times which means the protocol has the lowest capacity of receiving valid data packets. (c) The proposed fusion algorithm ACOA-AFSA can achieve much more data traffic with less energy consumption lies in such a fact; ACOA-AFSA algorithm makes use of biological intelligence to find global path for information delivery in minimum duration. What is more, the proposed algorithm can adaptively select the best route to transmit data if network topology changes. Given the reasons above, the proposed algorithm performs better than the other two algorithms.

As ACOA-AFSA algorithm has the best routing path for data transmission, number of total data packets received at sink node is the largest for certainty. VBF and LEACH perform worse than ACOA-AFSA algorithm. VBF protocol takes data forwarding as its most important task if monitoring network is large enough, which results in VBF protocol's higher network load. That is to say, utilization ratio of VBF protocol's network is pretty lower which leads to lower sink node's lower data reception. As clustering nodes are selected randomly in LEACH protocol, some selected clustering nodes are not fit for the mission as their energy is lower. In other words some clusters may lose their clustering nodes easily for the selected nodes' lower energy, which brings about sink node lower data reception. Figure 4 proves the proceeding discussion.

(b) Loss Ratio of Data Packets. Figure 5 shows VBF and LEACH have higher packets dropping rate even if nodes have sufficient energy for data receiving and transmitting. The main reasons for such phenomenon is that VBF and LEACH are easily affected by dynamic and complicated underwater environment. Besides in VBF protocol nodes only near routing vector can transmit data which means VBF protocol has lower network usage ratio. As a result loss rate of VBF is largest among three UWSN routing protocol. LEACH protocol allows clustering node to realize the data transmission while nonclustering nodes can restore their energy for the next data transmission. Thus LEACH protocol performs better than VBF protocols. For the same reason, clustering nodes, which is chosen randomly

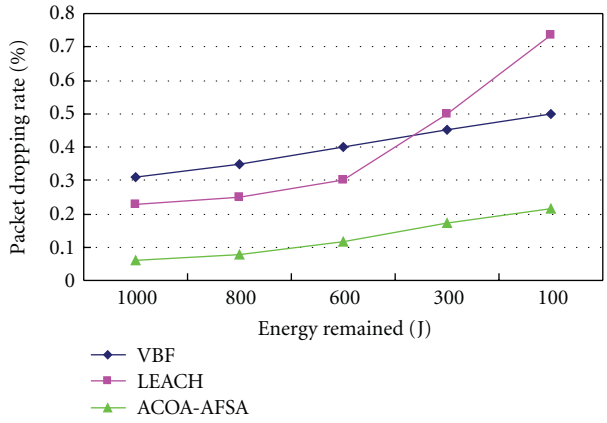


Figure 5: Packet dropping rate versus remaining energy.

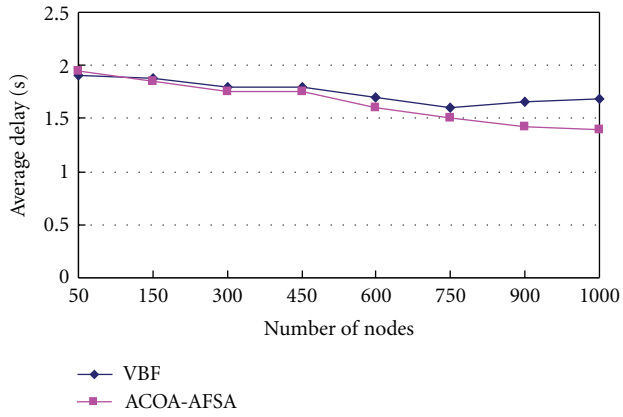


Figure 6: Average transmission delay versus number of nodes.

in LEACH protocol, are easy to fail as they execute the assignments which consume a lot of energy. After several data transmission circles, energy distribution is extremely unevenly distributed. So nodes' energy in some monitoring regions is super small that means data delivery in the area is likely to fail. Through our experiment, it can be revealed that ACOA-AFSA gets the least data dropping probability. Such an outperformance may exist owing to ACOA-AFSA's global and local view of the whole network topology. And simultaneously, the self-adaptive switching of ACOA and AFSA causes less energy consumption during the disruption of the clusters resulting in deterioration of the performances.

(c) Average Delay. We have also conducted the simulation to observe the average delay over number of nodes, namely, the relationship between network scale and average delay. VBF protocol is kind of UWSN routing protocol that performs well while scale of network is small. The major reason is that VBF protocol empowers most nodes in network to transfer data with less data retransmission if UWSN is dense network. As nodes' number increases, the protocol's shortage of data retransmission handicaps network's normal operation of data receiving and delivering seriously. Simulation result in Figure 6 testifies aforementioned VBF protocol's merits and demerits as well. What is more, Figure 6 demonstrates that ACOA-AFSA fusion algorithm could lower the average delay more with number of nodes increasing. The reason that ACOA-AFSA performs better than VBF is that the self-adaption mechanisms have been incorporated into the protocol. So, the routine efficiency is not so sensitive to the network scale and the scalability is thus optimized.

5. Conclusion

In this paper, a novel ACOA-AFSA fusion algorithm for UWSN routing protocol has been presented. It is a useful routing algorithm for underwater sensor networks owing to its local acknowledge and global view offered by ACOA and AFSA, respectively. The proposed algorithm also introduces a self-adaptive mechanism to fuse such two algorithms for searching better routing path. ACOA algorithm's parallel feature makes routing search in proposed protocol quickly. As AFSA algorithm finds candidate optimal routing path roughly in the fusion algorithm for optimizing and ACOA algorithm helps AFSA to calibrate routing path, fusion algorithm's robustness is stronger than VBF and LEACH. The fusion algorithm's global view and local optimal routing path makes its energy consumption for data transmission and reception more efficiently. With the increment of nodes and iteration number, advantages of swarm intelligence play greater role in finding best transmission route as well. The whole train of simulations proves ACOA-AFSA routing algorithm outperforms VBF and LEACH marginally in the way of energy consumption, packet loss rate, and delay. For future research work in the field, we will devote our efforts to the optimal tuning of parameters in ACOA and AFSA algorithm to realize the optimal performance in underwater wireless sensor network's data routing.

Acknowledgments

This paper was supported by Foundation of Shanghai Maritime University (20110028), Shanghai Science and Technology Committee Program (11PJ1404300, 09170502000), Shanghai Leading Academic Discipline Project (S30602), and a Grant from the National Natural Science Foundation of China (40801174).

References

[1] Z. Yanhua, *Research and Design of Routing Algorithm for Underwater Sensor Network*, Ocean University of China, Qingdao, China, 2010.

[2] Z. Jie, *The Study of Routing Protocols for Underwater Sensor Networks*, Huazhong University of Science & Technology, Wuhan, China, 2007.

[3] W. Changying, Z. Jie, W. Yanfeng, and G. Sheng, "Route s optimization and simulation of underwater sensor network VBF protocol," *Periodical of Ocean University of China*, no. 6, pp. 1261–1264, 2009.

[4] S. Guizhi and S. Enfang, "An energy-efficient routing protocol for underwater acoustic sensor networks," *Technical Acoustics*, no. 1, pp. 134–136, 2007.

[5] D. Haibin, *Ant Colony Optimization and Its Applications*, Sciences Press, Bejing, China, 2006.
[6] L. Xiaolei, *A New Intelligent Optimization Method-Artificial Fish Swarm Algorithm*, Zhe Jiang University, Hangzhou, China, 2003.
[7] W. Lianguo and S. Qiuhong, "Parameters analysis of artificial fish swarm algorithm," *Computer Engineering*, no. 24, pp. 169–171, 2010.
[8] C. Lv, S. Wang, and M. Tan, "Survey on mobile underwater wireless sensor networks," *Kongzhi yu Juece/Control and Decision*, vol. 24, no. 6, pp. 801–812, 2009.
[9] Z. Guo, H. Luo, F. Hong, M. Yang, and L. M. Ni, "Current progress and research issues in underwater sensor networks," *Computer Research and Development*, vol. 47, no. 3, pp. 377–389, 2010.
[10] Y. Xiao, Y. Zhang, J. H. Gibson, G. G. Xie, and H. Chen, "Performance analysis of ALOHA and p-persistent ALOHA for multi-hop underwater acoustic sensor networks," *Cluster Computing*, vol. 14, no. 1, pp. 65–80, 2011.
[11] S. Qiliang and Z. Liping, "ACO-based routing algorithm for wireless sensor network," *Computer Measurement & Control*, no. 5, pp. 1253–1256, 2011.
[12] C. Xuhua, Z. Zhijin, and Y. Xueyi, "Network route reconfiguration based on improved ant algorithm," *Computer Engineering*, no. 13, pp. 90–92, 2009.
[13] Y. Xiangguang, Z. Yongquan, and L. Yongmei, "Hybrid algorithm with artificial fish s warm algorithm and PSO," *Application Research of Computers*, no. 6, pp. 2084–2087, 2010.
[14] G. Mingjia, X. Shibin, L. Kanchao, and L. Yongsheng, "QoS routing algorithm based on combination of modified ant colony algorithm and artif icial fish sw arm algorithm," *Computer Technology and Development*, no. 7, pp. 145–148, 2009.
[15] M. Ayaz and A. Abdullah, "Hop-by-hop dynamic addressing based (H2-DAB) routing protocol for underwater wireless sensor networks," in *International Conference on Information and Multimedia Technology (ICIMT '09)*, pp. 436–441, December 2009.
[16] M. C. Domingo and R. Prior, "Design and analysis of a GPS-free routing protocol for underwater wireless sensor networks in deep water," in *International Conference on Sensor Technologies and Applications (SENSORCOMM '07)*, pp. 215–220, October 2007.
[17] C. J. Huang, Y. W. Wang, C. F. Lin, Y. T. Chen, H. M. Chen, and H. Y. Shen, "A self-healing clustering algorithm for underwater sensor networks," *Cluster Computing*, no. 1, pp. 91–99, 2011.
[18] J. Chengzhi, *Artificial Intelligent Technique*, Tsinghua University Press, Bejing, China.
[19] Z. Huijun, Z. Wei, and T. Shaohua, "Improved ACO-based wireless sensor networks routing," *Application Research of Computers*, no. 1, pp. 99–101, 2010.
[20] F. Yuexi, J. Xinyu, and C. Wenyu, "Wireless sensor network routing protocol based on improved ant colony algorithm," *Chinese Journal of Sensors and Actuators*, no. 11, pp. 2461–2464, 2007.
[21] Z. Jian, H. BenXiong, Z. Fan, and T. Lai'an, "Energy efficient routing protocol applied in underwater sensor networks," *Computer Science*, vol. 35, pp. 38–42, 2008.
[22] X. Hong-Mei, C. Yi-Bao, L. Jia-Guang, and W. Yan-Tao, "The research on the parameters of the ant colony algorithm," *Journal of Shandong University of Technology (Natural Science Edition)*, no. 1, pp. 7–11, 2008.
[23] J. Lingyan, Z. Jun, and Z. Shuhong, "The analysis on the parameters of ACOA," *Computer Engineering and Applications*, no. 20, pp. 31–36, 2007.
[24] X. Li, S. L. Fang, and Y. C. Zhang, "The study on clustering algorithm of the underwater acoustic sensor networks," in *14th International Conference on Mechatronics and Machine Vision in Practice (M2VIP '07)*, pp. 78–81, December 2007.
[25] P. Xie, J. H. Cui, and L. Lao, "VBF: vector-based forwarding protocol for underwater sensor networks," in *5th International IFIP-TC6 Networking Conference, Networking*, vol. 3976 of *Lecture Notes in Computer Science*, pp. 1216–1221, 2006.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 542389, 9 pages
doi:10.1155/2012/542389

Research Article

A Local World Evolving Model for Energy-Constrained Wireless Sensor Networks

Nan Jiang, Huan Chen, and Xiang Xiao

College of Information Engineering, East China Jiaotong University, Nanchang 150001, China

Correspondence should be addressed to Nan Jiang, jiangnan1018@gmail.com

Received 23 October 2011; Accepted 23 February 2012

Academic Editor: Sabah Mohammed

Energy efficiency is one of the basic requirements of wireless sensor networks (WSNs) yet this problem has not been sufficiently explored in the context of cluster-based sensor networks. Specifically, the interaction of different types of sensor nodes is one of the major factors of energy efficiency in large-scale heterogeneous networks. In this paper, we aim at improving the interactions among sensor nodes, and we present a heterogeneous local-world model to form large-scale wireless sensor networks based on complex network theory. Two types of nodes, normal nodes and cluster nodes, are added to the networks. The degree distribution for this model is obtained analytically by mean-field theory. This approach depicts the evolution of the network as having a topological feature that is not completely exponential and not completely power law; instead, it behaves between them. The experiment and simulation indicate that the control method has excellent robustness and a satisfactory control effect (interaction of different types of sensor nodes). Furthermore, the results also show that the higher the generation rate of the cluster nodes is, the closer the degree distribution follows the power-law distribution.

1. Introduction

Wireless sensor networks (WSNs) are composed of a large number of sensor nodes. These nodes are deployed in a specific area to monitor physical phenomena. Wireless sensor networks are poised to revolutionize our abilities for sensing and controlling our environment. However, sensor nodes rely on batteries to provide energy in WSNs. As a result, the primary problem in WSNs is always energy savings for sending and receiving data. Consequently, power control must be performed in WSNs. In recent years, with the development of WSNs, advances in wireless communication and sensing technology have made it possible to produce a large number of individually cheap and small units. Usually, these small pieces of equipment are used in WSNs, as combinations of sensors and wireless network nodes. With the increasing scale of WSNs, some characteristics of complex networks have emerged. The problem of coefficient, error, and attack tolerance remains to be solved in large-scale networks. As a result, it is crucial to find out a good way to handle these problems.

In recent years, complex networks have received increasing attention for exhibiting the topological structure, function and dynamical properties of many real-world networks such as the World Wide Web, the Internet [1, 2], social networks [3], biological networks [4], and ad hoc networks [5]. One of the most important discoveries is the B-A model [6]. This model is based on two foundational mechanisms: growth and preferential attachment. A new node is added to the network at each step and connects with an old node with a specific probability, which is related to the degree of the old node. The B-A network has the scale-free property and follows the power-law distribution.

While the B-A scale-free network model captures the basic mechanism that is responsible for the power-law degree distribution, it is still a minimal model with several limitations. As stated in [7], Li Xiang and Chen Guanrong proposed a Li-Chen model to modify a limitation in the B-A model. They suggested that there should be a local world of each node in various real-world complex networks. Moreover, the preferential attachment mechanism of a scale-free network does not work in a global network but does work in the local world of each node. The Li-Chen model represents a transition between a power-law and an exponential distribution. The BA scale-free model is only one of its

special cases. This case fits the real circumstance, in which new nodes enter the networks, and a new node chooses only some of its local world nodes to link with. Numerical results have shown that networks constructed with a local preferential attachment mechanism can maintain the robustness of scale-free networks undergoing random errors and can concurrently improve the reliance against targeted attacks on highly connected nodes [6, 8].

2. Related Work

Because cluster structure networks can be modified to obey the property of scale-free, many people attempt to adjust the traditional structure of WSNs to a new structure of scale-free. Thus, sensor nodes can save energy even if there is a large-scale network in the monitoring area. They obtain a satisfactory result when forming WSNs that possess the scale-free character as well. It is shown that scale-free networks are robust against the random removal of nodes or node failures. Therefore, another advantage of this type of WSN is the robustness of the network, which is a groundbreaking result in this research field and provides good inspiration for this paper.

In 2007, Chen et al. [9] studied a new evolving mechanism for deducing the fault-tolerant communication topology among the cluster heads with complex network theory. Based on the B-A model's growth and preferential attachment element, they not only used a local-world strategy for the network when a new node was added to its local-world but also selected a fixed number of cluster heads in the local world, for the purpose of obtaining a good performance in terms of random error tolerance. In 2009, Zhu et al. adopted two evolving algorithms for WSNs, which are based on complex network theory [10]. The first model was applied to deduce energy-aware communication topology, and this model can generate scale-free networks that have satisfactory performance of random error tolerance. The second model involved limiting the number of communication links for each node base in the first model. Moreover, based on the model in Hailin Zhu's paper, Li et al. investigated an improved local-world model of WSNs in 2011 [11]. There are two types of nodes in [11]: normal nodes and cluster nodes, which is different from the homogeneous WSNs. The authors created a heterogeneous evolving network model. In addition, they limited the number of linking edges in every cluster node to form a good scale-free property in WSNs and to control the energy consumption in the sensor nodes. However, they did not consider the tendency for interactions between different types of sensor nodes, which could reduce the energy efficiency (to a variable degree) in WSNs.

In this paper, an evolving model for WSNs based on the B-A model and the Li-Chen model [7] is proposed. Aimed at improving the interaction among different types of sensor nodes in WSNs, the proportion of a normal node in a cluster node's neighbor nodes is increased in this model.

The remainder of this paper is organized as follows. In Section 3, we propose an algorithm for an energy-aware evolution model. In Section 4, we analyze the model by mean-field theory, and we provide numerical experiments that demonstrate the features of the networks that are generated by the proposed algorithms. Finally, Section 5 gives the conclusions of this paper.

3. The Generalized Local-World Model

3.1. The Local-World Model. A local world is a small community with a few nodes in a large-scale network. The model of the local-world is deduced from the B-A model. In the Li-Chen model, each node has only local connection information; nodes connect only in their local world based on local connection information. Then, M nodes are selected randomly from the existing network as the "local world" of the newly added node. The network generation algorithm of the Li-Chen model can be described in two steps, as follows

(1) Growth. Starting from a small number m_0 of isolated nodes, at each time step t, add a new node with m edges connecting to the network.

(2) Local Preferential Attachment. At each time step t, before connecting the new incoming node to m existing nodes, randomly select M nodes as its local-world; then, connect m nodes in its local world, using a local preferential attachment with probability $\prod_{\text{Local}}(k_i)$, which is determined at every time step t by the following:

$$\prod_{\text{Local}}(k_i) = \prod{}'(i \in \text{Local-world}) \frac{k_i}{\sum_j \text{Local}\, k_j}, \tag{1}$$

where $\prod'(i \in \text{Local-world}) = M/(m_0 + t)$.

After t time steps, there will be a network with $N = t + m_0$ nodes and $m * t$ edges.

3.2. Generalized Local-World Models for Wireless Sensor Networks. In this section, we model a WSN as an inhomogeneous network with growth and preferential attachment. This model contains two types of nodes: normal nodes and cluster nodes. There is only one cluster node attached to a normal node; in other words, the normal node has only one edge, which means that the normal node cannot relay data from other nodes. A cluster node can integrate and transmit data from other nodes. Both of the two types of nodes can connect to a cluster node, and the number of edges is limited in every cluster node because of its energy efficiency. As every new cluster node joins the network, it is randomly assigned an initial energy E_i from [0.5, 1] to the cluster node, while every normal node is set to 0. The limited number of edges in every cluster node is represented by $k_{\max i}$, which is based on the initial energy of the cluster nodes E_i, and [0.5, 1] is chosen as the interval $[E_{\min}, E_{\max}]$. We suppose here that $k_{\max i}$ [10] is given as follows:

$$k_{\max i} = k_{\max} \frac{E_i}{E_{\max}}, \tag{2}$$

where $k_{\max i}$ reflects the ability of having the maximum number of edges for cluster node i.

Then, the growing model is described as follows: starting with a small number of nodes, m_0 (all of them are cluster nodes), they randomly link each other. The result constitutes the initial network.

(1) Growth. At every time step, a new cluster node or a normal node with one edge enters into the existing network with a probability p or $1 - p$, respectively. If the new node is a cluster node, then give it a random energy value E_i. The probability p is the proportion of cluster nodes in the evolving model. A small number of cluster nodes would cause many sensor nodes to link to them, which results in working frequently and consuming energy fast; however, a large number of cluster nodes would waste resources and decrease the energy efficiency. Thus, we assume $0 < p \leq 0.5$.

(2) Preferential Attachment. The new incoming node links to an old cluster node that is selected randomly from the preexisting network. Nodes in WSNs have the constraint of energy and connectivity and only communicate data with the cluster nodes in their local area. First, M cluster nodes are selected randomly from the network as the new incoming node's local world $\boldsymbol{\Omega}$; then, one of the cluster nodes is chosen to link with the new node according to the probability $\prod_i$.

(1) If the new incoming node is a cluster node, then the probability is set as follows:

$$\prod_{(k_i)} = \left(1 - \frac{k_i}{k_{\max i}}\right) \frac{k_i}{\sum_{j \in \text{local}} k_j}. \tag{3}$$

In this case, when the value of k_i is high, the probability that it will be chosen to connect with the new node is higher.

(2) If the new incoming node is a normal node, then the probability is adapted as follows:

$$\prod_{(k_i)} = \left(1 - \frac{k_i}{k_{\max i}}\right) \frac{s_i}{\sum_{j \in \text{local}} s_j}, \tag{4}$$

where s_i is the number of cluster node i's edges that should connect with the cluster node on both sides, and the greater the value of s_i is, the higher the probability that it will be chosen to connect with the new node. Only through this approach we can adjust the number of cluster nodes that are linked to one cluster node (cluster head).

In [10], the authors consider the expenditure of energy in the process of linking nodes together. The disadvantage is that the energy in a cluster node would be exhausted in only several rounds. In fact, the energy consumption is relatively low; thus, the energy consumption is not considered in this stage, and only $k_{\max i}$ is considered to be the limit for a cluster node to connect to others randomly.

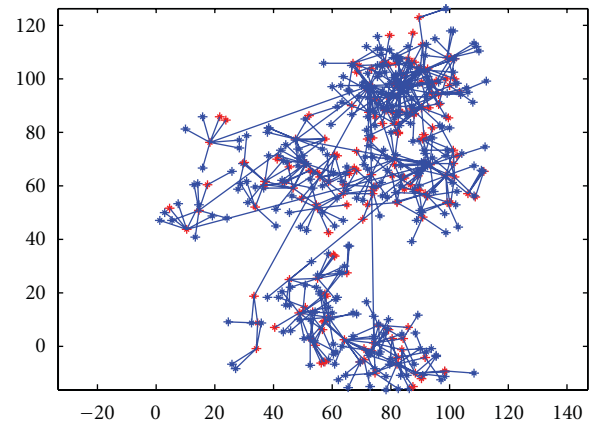


Figure 1: The scenario of forming WSNs over 500 times when $p = 0.25$, $m_0 = 10$, and $M = 4$. The red dots stand for cluster nodes, and the other dots are normal nodes. We can observe that there are several main unequal clusters in the graph, and they form basic wireless sensor networks.

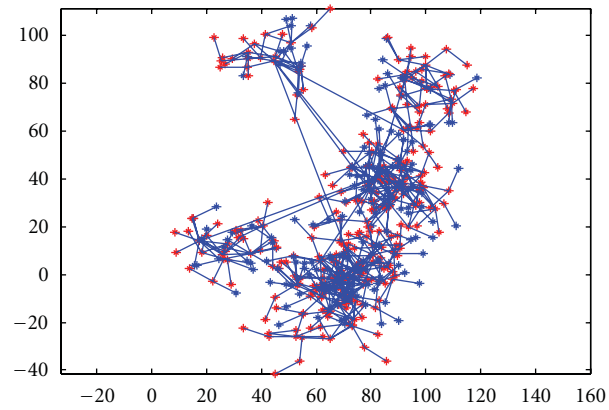


Figure 2: The scenario of forming WSNs over 500 times when $p = 0.5$, $m_0 = 10$ and $M = 4$.

(3) After step (2), if the degree $k_i \geq k_{\max i}$, then return to step (2).

From $(1 - k_i/k_{\max i})$, we can know that the closer k_i is to $k_{\max i}$, the lower the probability that the node would be connected to the new incoming nodes. When k_i is closer to its maximum degree $k_{\max i}$, node i will not connect to the new incoming nodes any more.

From the discussion above, the whole probability $\prod_i$ is as follows:

$$\prod_i = p \frac{k_i}{k_i / \sum_{j \in \text{local}} k_j} + (1 - p) \frac{s_i}{s_i / \sum_{j \in \text{local}} s_j}. \tag{5}$$

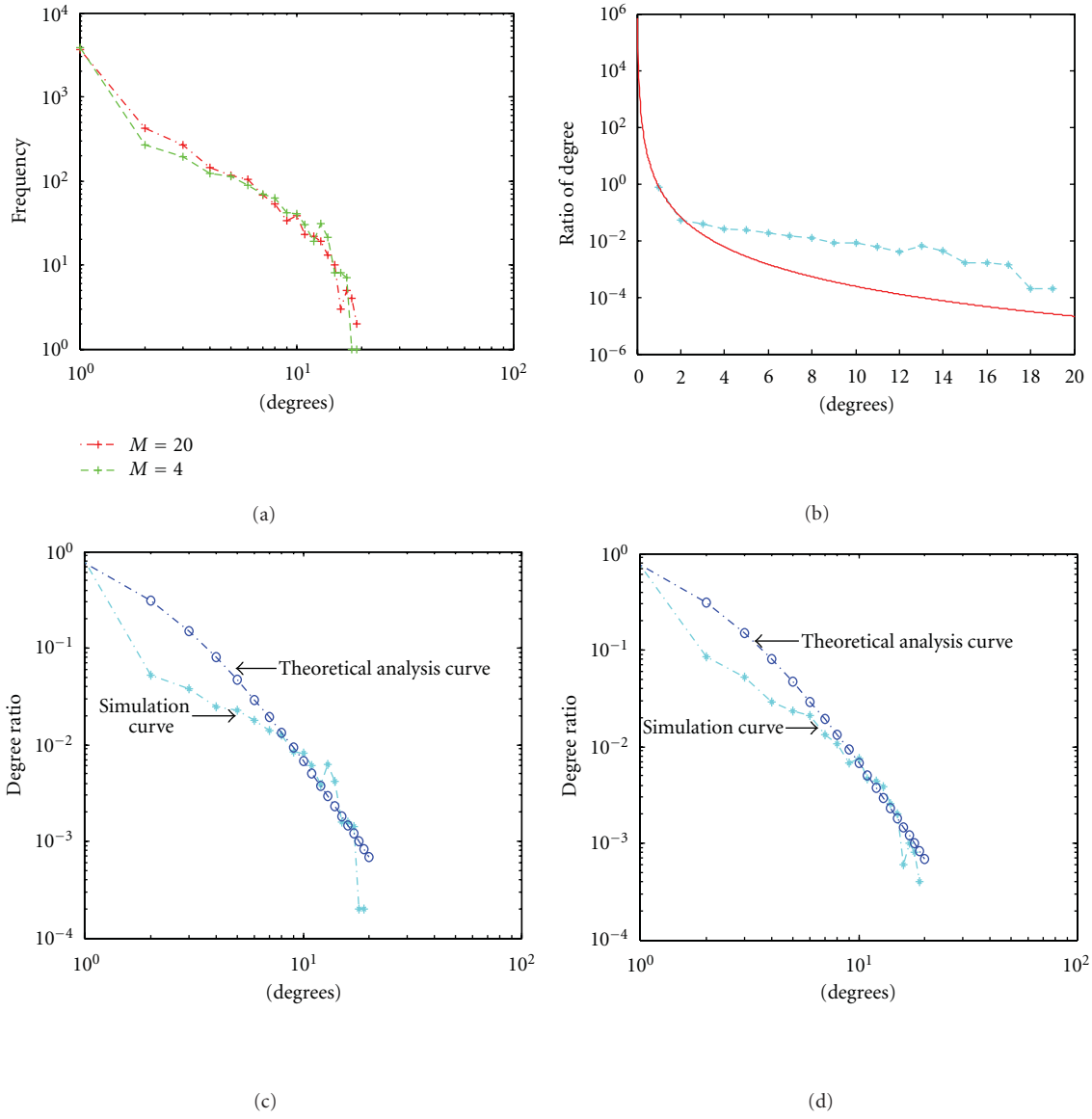


FIGURE 3: We set our simulation parameters to be $p = 0.25$ and $N = 5000$ in Scenario1. (a) The degree distribution plotted on a log-log scale of the proposed model. The first curve is for $m_0 = 20$, $M = 20$, and the second curve is for $m_0 = 10$, $M = 4$. (b) The plot of the simulation result for $m_0 = 10$, $M = 4$. The solid line is the fitted curve from the simulation result. (c) Degree distribution comparison on a log-log scale, comparing the proposed model with theoretical analysis from (24) when $m_0 = 10$, $M = 4$. (d) Degree distribution comparison on a log-log scale, comparing the proposed model with theoretical analysis from (24) when $m_0 = 20$, $M = 20$.

4. Analytical Results of the Model for WSNs

In this section, using mean-field theory [6], the theoretical analysis of the degree distribution $p(k)$ of the nodes is given in the model:

$$\begin{aligned}\frac{\partial k_i}{\partial t} &= \frac{M}{m_0 + pt}\prod(i) \\ &= \frac{M}{m_0 + pt} \\ &\quad \times \left[p\left(1 - \frac{k_i}{k_{\max i}}\right)\frac{k_i}{\sum_j k_j} \right. \\ &\qquad \left. + (1-p)\left(1 - \frac{k_i}{k_{\max i}}\right)\frac{s_i}{\sum_j s_j} \right].\end{aligned} \tag{6}$$

Case 1 ($M = 1$). When $M = 1$, the new node selects only one node in the local world $\boldsymbol{\Omega}$ unless it reaches k_i. Moreover, the preferential attachment mechanism does not work. Thus, the growth rate of k_i is as follows:

$$\frac{\partial k_i}{\partial t} = \frac{1}{m_0 + pt}. \tag{7}$$

The denominator of the above equation is the number of cluster nodes in time t. From the initial condition $k_i = 1$, we can solve the differential equation and find that the degree

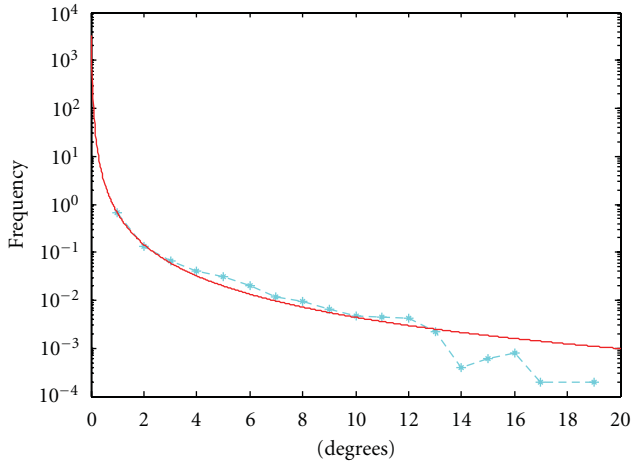


FIGURE 4: Plot of the simulation result for $p = 0.5$, $m_0 = 10$ and $M = 4$. The solid line is the fitted curve from the simulation result.

distribution follows an exponential distribution in spite of the constraint $k < k_{\max}$. Thus, we have the following:

$$p(k) \propto \exp(-pk)(t \longrightarrow \infty). \tag{8}$$

Case 2 ($M = m_0 + pt$). This case means that the local world $\mathbf{\Omega}$ is the whole network; thus, (6) will become the following:

$$\begin{aligned}\frac{\partial k_i}{\partial t} &= \frac{M}{m_0 + pt}\prod(i)\\ &= p\left(1 - \frac{k_i}{k_{\max i}}\right)k_i + (1-p)\left(1 - \frac{k_i}{k_{\max i}}\right)s_i.\end{aligned} \tag{9}$$

In a network, the degrees k_i of most of the nodes are much smaller than their maximum $k_{\max i}$; thus, similarly, we obtain the following:

$$1 - \frac{k_i}{k_{\max i}} \approx 1. \tag{10}$$

Then, we can simplify (9) as the following:

$$\begin{aligned}\frac{\partial k_i}{\partial t} &= p\left(1 - \frac{k_i}{k_{\max i}}\right)k_i + (1-p)\left(1 - \frac{k_i}{k_{\max i}}\right)s_i\\ &= p\frac{k_i}{\sum_j k_j} + (1-p)\frac{s_i}{\sum_j s_j}.\end{aligned} \tag{11}$$

To calculate $\prod(i)$, we first need to compute the parameters of k and s:

$$\bar{k} = \frac{\sum_j k_j}{m_0 + pt} = \frac{2(m_0 + t) - (1-p)t}{m_0 + pt} = \frac{2m_0 + t + pt}{m_0 + pt}, \tag{12}$$

$$\bar{s} = \frac{\sum_j s_j}{m_0 + pt} = \bar{k} - \frac{(1-p)t}{m_0 + pt} = 2. \tag{13}$$

From (12) and (13), we know the following:

$$\begin{aligned}\frac{\partial k_i}{\partial t} &= p\frac{k_i}{\sum_j k_j} + (1-p)\frac{s_i}{\sum_j s_j}\\ &= p\frac{k_i}{(m_0 + pt)\,\bar{k}} + (1-p)\frac{s_i}{(m_0 + pt)\,\bar{s}}\\ &= p\frac{k_i}{2m_0 + pt + t} + (1-p)\frac{s_i}{(m_0 + pt)2}.\end{aligned} \tag{14}$$

When $t \to \infty$, we can obtain the following:

$$\frac{\partial k_i}{\partial t} = p\frac{k_i}{pt + t} + (1-p)\frac{s_i}{2pt}. \tag{15}$$

Case 3 ($1 < M < m_0 + pt$). In this case, a newly added node selects one of the cluster nodes in its local world $\mathbf{\Omega}$, and we can obtain the following:

$$\begin{aligned}\frac{\partial k_i}{\partial t} &= \frac{M}{m_0 + pt}\prod(i)\\ &= \frac{M}{m_0 + pt}\left[p\frac{k_i}{\sum_j k_j} + (1-p)\frac{s_i}{\sum_j s_j}\right]\\ &= \frac{M}{m_0 + pt}\\ &\quad\times\left[p\frac{k_i}{M((2m_0 + pt + t)/(m_0 + pt))} + (1-p)\frac{s_i}{2M}\right]\\ &= \frac{p}{2m_0 + pt + t}k_i + (1-p)\frac{s_i}{2(m_0 + pt)}\\ &\approx \frac{p}{pt + t}k_i + (1-p)\frac{s_i}{2pt}.\end{aligned} \tag{16}$$

A comparison of (15) with (16) shows that they have almost the same solution under the same simple assumption. To calculate the above equation, we assume here that $s_i = 2$. Thus, (16) can be written as the following:

$$\begin{aligned}\frac{\partial k_i}{\partial t} &= \frac{p}{2m_0 + pt + t}k_i + (1-p)\frac{s_i}{2(m_0 + pt)}\\ &= \frac{p}{pt + t}k_i + \frac{1-p}{pt}\\ &= \frac{p}{p+1}\frac{k_i}{t} + \frac{1-p}{p}\frac{1}{t}.\end{aligned} \tag{17}$$

We define $A = p/(p+1)$ and $B = (1-p)/p$, and we can obtain the changeable equation (16):

$$\frac{\partial k_i}{\partial t} = A\frac{k_i}{t} + B\frac{1}{t}. \tag{18}$$

Then, we can obtain the following:

$$\frac{\partial k_i}{Ak_i + B} = \frac{\partial t}{t}. \tag{19}$$

Combining the initial condition $k_i(t_i) = 1$, we can obtain the solution of the previous equation:

$$k_i(t) = \left(\frac{t}{t_i}\right)^A \frac{A+B}{A} - \frac{B}{A}. \tag{20}$$

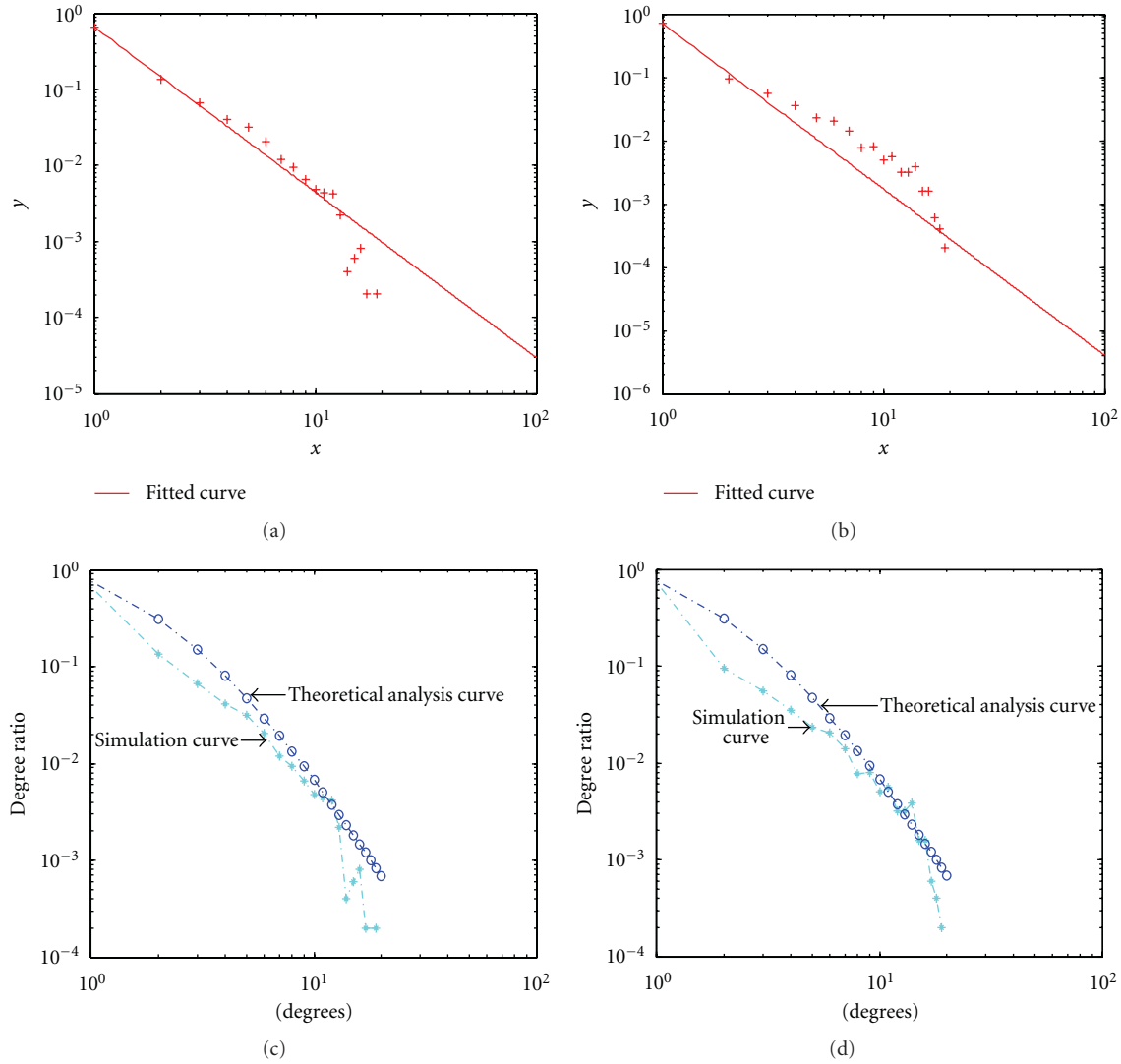


FIGURE 5: We set our simulation parameters to be $p = 0.5$ and $N = 5000$ in Scenario2. (a) Plot of the simulation result for $m_0 = 10$, $M = 4$. The solid line is the fitted curve from the simulation result in log-log coordinates. (b) The plot of the simulation result for $m_0 = 20$, $M = 20$. The solid line is the fitted curve from the simulation result in log-log coordinates. (c) Degree distribution comparison on a log-log scale, comparing the proposed model with theoretical analysis from (24) when $m_0 = 10$, $M = 4$. (d) Degree distribution comparison on a log-log scale, comparing the proposed model with theoretical analysis from (24) when $m_0 = 20$, $M = 20$.

Moreover, to find the degree distribution $P(k)$ (i.e., the probability that a node has k edges), we first calculate the cumulative probability $P[k_i(t) < k]$. Suppose that the node enters into the network in equal time intervals. We define the probability density of t_i as follows:

$$P(t_i) = \frac{1}{m_0 + t}. \tag{21}$$

Then, $P[k_i(t) < k]$ has the following form:

$$\begin{aligned} P[k_i(t) < k] &= P\left(t_i > t\left(\frac{A+B}{Ak+B}\right)^{1/A}\right) \\ &= 1 - \frac{t}{t+m_0}\left(\frac{A+B}{Ak+B}\right)^{1/A}. \end{aligned} \tag{22}$$

Hence, the degree distribution $P(k)$ can be obtained:

$$\begin{aligned} P(k) &= \frac{\partial P[k_i(t) < k]}{\partial k} \\ &= \frac{t}{m_0 + t}(A+B)^{1/A}(Ak+B)^{-1/A-1} \\ &\approx (A+B)^{1/A}(Ak+B)^{-1/A-1}. \end{aligned} \tag{23}$$

Next, inserting A and B back into $P(k)$, we obtain the final equation of $P(k)$:

$$P(k) = \left(\frac{1}{p+p^2}\right)^{1+1/p}\left(\frac{p}{1+p}k + \frac{1-p}{p}\right)^{-2-1/p}. \tag{24}$$

To compare different ratios of cluster nodes to the whole network's forming factor, in this case, we take two different

ratios of p to show the diversity of our simulations. In addition, we make another comparison by changing M in the case of different ratios p. Here, all of the simulations are operated by MATLAB, which are based on our simulation toolboxes. The two types of p are 0.25 and 0.5.

When we set $p = 0.25$, the emergence of a cluster node is much less than a normal node. As we show in Figure 1, the sensor nodes have a highly concentrated degree. Furthermore, most nodes join the largest group; this phenomenon can be explained by the theory of the B-A scale-free network model. A newly generated node joins the cluster node that contains a higher number of edges; in the long run, this strategy can lead to the phenomenon of the Matthew effect, in which a small number of nodes have the most edges and the remaining nodes have the least edges. On the other hand, when we change $p = 0.5$, the emergence of the cluster nodes is equal to the normal nodes. As depicted in Figure 2, the sensor nodes have a highly concentrated degree and join the largest group.

The degree distribution does not follow the power-law distribution but instead distributes between a power-law and an exponential distribution, and the frequency of the degree in the whole network is shown in Figure 3(a). In Figure 3(b), the picture shows the relationship between the degree and its ratio. Another standard line (the fitted straight line) is drawn from the data to contrast with the ratio line. The figure shows that the two lines do not fit well. In another words, in this situation, the degree distribution does not follow a scale-free distribution. Figure 3(c) and Figure 3(d) compare the experiment data with the theoretical model in (24).

The relationship between the degree and its ratio $p = 0.5$ is shown in Figure 4. We can observe that the fitted line fits the experiment data perfectly; this finding indicates that, in this case, the edge distribution gets close to the power-law distribution. Furthermore, the two lines are close to the line of the power-law distribution in Figures 5(a) and 5(b).

Figures 5(c) and 5(d) not only show the ratio of the difference degree in the actual condition and the theoretical condition but also indicate that the two lines match closely with each other. As analyzed above, the model that we proposed has some restriction on the growth of a scale-free network, with the result that the simulation result is different from the theoretical model.

Because the distribution of the node degree could not uncover the internal complexity of the topological structure, it is worth investigating the connectivity correlation among the cluster nodes, which could lead to a more profound understanding of the structure. For the sake of depicting the relationship among the cluster nodes and normal nodes, we first discuss the property of the average nearest-neighbor connectivity. Because the cluster nodes play the role of a backbone in ensuring the data processing performance of the WSNs, we compute this performance here as in [12]:

$$k_{nn}(k) = \frac{1}{l_k} \sum_{i \in M} \left(\frac{1}{k} \sum_{j \in l_i} a_{ij} l_j \right), \tag{25}$$

where l_k is the number of cluster nodes with degree k in the WSNs and l_k is the neighbor node set of a cluster node i with

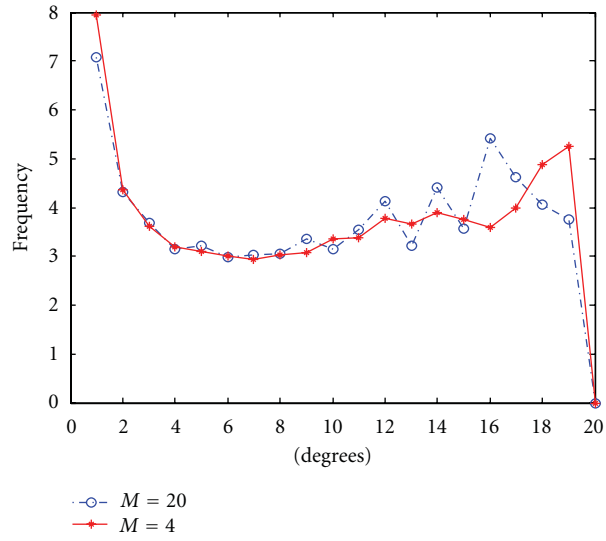


Figure 6: Degree distribution and its average frequency in neighboring nodes.

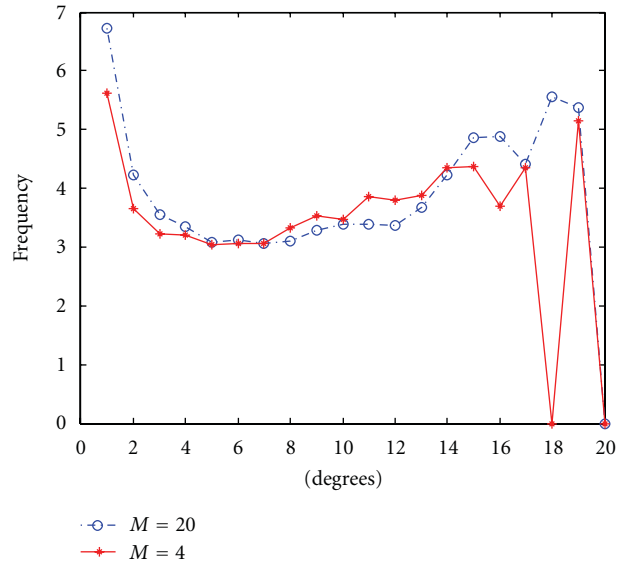


Figure 7: Degree distribution and its average frequency in neighboring nodes when $p = 0.5$, $N = 5000$. The first curve is for $m_0 = 20$, $M = 20$, and the second curve is for $m_0 = 10$, $M = 4$.

degree k. Additionally a_{ij} is an element of the adjacent matrix of the cluster nodes, and k_j is the degree of a cluster node j. Figures 6 and 7 show the $k_{nn}(k)$ curve for different values of M. We can observe in both pictures that the frequency of most of the degrees is approximately 3 to 4.

Second, to show our link mechanism's property, Figures 8(a) and 8(b) are adopted to describe the result of the interaction among different types of sensor nodes between the two types of nodes. Figures 8(a) and 9(a) show the edge number of every cluster node linking to the neighbor cluster node, and most of the cluster nodes distribute in the frequency

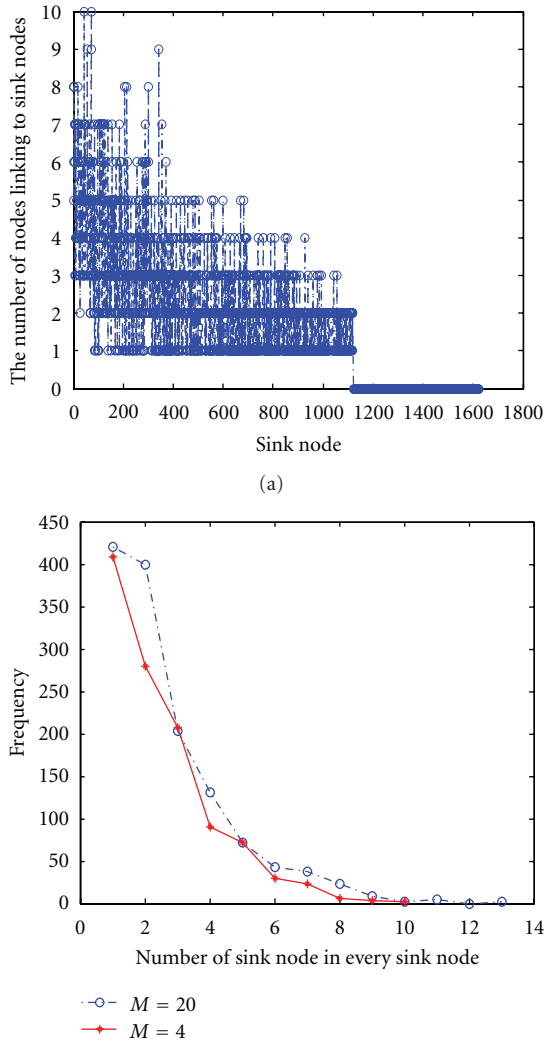


Figure 8: We set our simulation parameters to be $p = 0.25$ and $N = 5000$ in Scenario1. (a) The distribution of every cluster node's neighboring cluster nodes. Cluster node i is more likely to possess two cluster nodes in its neighbor node when $m_0 = 10$, $M = 4$. (b) The statistics of neighboring cluster nodes and their frequency. The first curve is for $m_0 = 20$, $M = 20$, and the second curve is for $m_0 = 10$, $M = 4$.

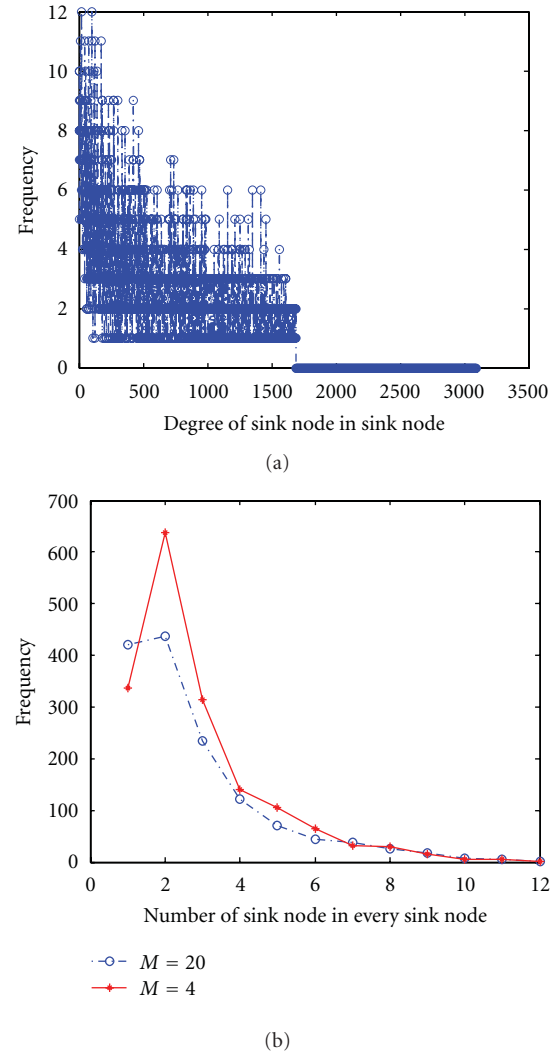


Figure 9: We set our simulation parameters to be $p = 0.5$ and $N = 5000$ in Scenario2. (a) The distribution of every cluster node's neighboring cluster nodes when $m_0 = 10$, $M = 4$. (b) The statistics of neighboring cluster nodes and their frequency. The first curve is for $m_0 = 20$, $M = 20$, and the second curve is for $m_0 = 10$, $M = 4$.

of 1 and 2. Moreover, Figures 8(b) and 9(b) depict the detailed distribution of cluster nodes for different values of M. A frequency of 2 is ranked in the second place from Figure 8(b), and when all of the data are calculated, a frequency of 2 is the average edge number of neighbor cluster nodes, which is also illustrated in Figure 9(b). The frequency of 2 is located in the second place, and we can find that the frequency of 2 is the mean of the edge number of neighbor cluster nodes, which is obtained from calculating the data. Thus, the conclusion of (13) is proven to be correct.

We can draw a conclusion that, with an increase of p (the proportion of cluster nodes), the degree distribution changes from an exponential distribution to a power-law distribution, which our experiments have proven from the discussion above.

5. Conclusions

This paper presents an evolving model for improving the energy efficiency of WSNs. This model adapts and balances the interactions of different types of sensor nodes to reach the goal of energy efficiency with two different types of probability in the evolving network. Furthermore, the proportion of edges linking to a normal node is improved in every cluster node. Our numerical experiments above show that our new model reaches the requirements for and improves on the network robustness. A good model adapts the cluster structure when the number of nodes increases continuously. However, it ignores the dead sensor nodes with the growth of the network in our model. For the next step, we will pay

more attention to the random deletion of sensor nodes that are too close to a real environment of WSNs.

Acknowledgments

This work is supported by National Natural Science Foundation of China under Grant nos. 61063037 and 61163056, by Key Projects in the Science & Technology Pillar Program of Jiangxi Province of China under Grant no. 20111BBG70031-2, and by Innovation Special Funds Projects for Graduate Students of Jiangxi Province under Grant no. YC2011-S079.

References

[1] L. A. Adamic and B. A. Huberman, "Growth dynamics of the world-wide web," *Nature*, vol. 401, no. 6749, p. 131, 1999.

[2] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Reviews of Modern Physics*, vol. 74, no. 1, pp. 47–97, 2002.

[3] J. Stehlé, A. Barrat, and G. Bianconi, "Dynamical and bursty interactions in social networks," *Physical Review E*, vol. 81, no. 3, pp. 1–4, 2010.

[4] H. Jeong, S. P. Mason, A. L. Barabási, and Z. N. Oltvai, "Lethality and centrality in protein networks," *Nature*, vol. 411, no. 6833, pp. 41–42, 2001.

[5] N. Sarshar and V. Roychowdhury, "Scale-free and stable structures in complex ad hoc networks," *Physical Review E*, vol. 69, no. 2, Article ID 026101, 6 pages, 2004.

[6] A. L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.

[7] X. Li and G. R. Chen, "A local-world evolving network model," *Physica A*, vol. 328, no. 1-2, pp. 274–286, 2003.

[8] S. W. Sun, Z. X. Liu, Z. Q. Chen, and Z. Z. Yuan, "Error and attack tolerance of evolving networks with local preferential attachment," *Physica A*, vol. 373, no. 1, pp. 851–860, 2007.

[9] L. J. Chen, D. X. Chen, X. Li, and J. N. Cao, "Evolution of wireless sensor network," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '07)*, vol. 556, pp. 3005–3009, March 2007.

[10] H. L. Zhu, H. Luo, H. P. Peng, L. X. Li, and Q. Luo, "Complex networks-based energy-efficient evolution model for wireless sensor networks," *Chaos, Solitons and Fractals*, vol. 41, no. 4, pp. 1828–1835, 2009.

[11] S. D. Li, L. X. Li, and Y. X. Yang, "A local-world heterogeneous model of wireless sensor networks with node and link diversity," *Physica A*, vol. 390, no. 6, pp. 1182–1191, 2011.

[12] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang, "Complex networks: structure, dynamics," *Physics Reports*, vol. 424, no. 4-5, pp. 175–308, 2006.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 820716, 13 pages
doi:10.1155/2012/820716

Research Article

A Wireless Sensor Network for Precise Volatile Organic Compound Monitoring

Gianfranco Manes,[1] Giovanni Collodi,[1] Rosanna Fusco,[2] Leonardo Gelpi,[2] and Antonio Manes[3]

[1] *University of Florence and The MIDRA Consortium, 50139 Florence, Italy*
[2] *Health, Safety, Environment and Quality Department, Eni S.p.A., 00144 Rome, Italy*
[3] *Netsens s.r.l, Sesto Fiorentino, 50019 Florence, Italy*

Correspondence should be addressed to Gianfranco Manes, gianfranco.manes@unifi.it

Received 24 November 2011; Accepted 8 February 2012

Academic Editor: Carlos Ramos

A variety of methods have been developed to monitor VOC concentration in hazardous sites. The methods range from calculation to measurement, point measuring to remote sensing. Some are suited for leak detection, others for estimation of the annual emission or both. None of the following available methods comes close to the ideal method. A distributed instrument providing precise monitoring of Volatile Organic Compound (VOC) concentration in a petrochemical plant is described; it consists of a Wireless Sensor Network (WSN) platform whose nodes are equipped with meteorological/climatic sensors and VOC detectors. Internet connectivity is provided in real time at a one-minute sampling rate, thus providing environmental authorities and plant management with an unprecedented tool for immediate warning in case of critical events. The paper describes the WSN platform, detailing various units (gateways, nodes, detectors) and shows the features of scalability and reconfigurability, with minimal intrusiveness or obtrusiveness. Environmental and process data are forwarded to a remote server and made available to the authenticated users through a rich user interface that provides data rendering in various formats and worldwide access to data. A survey of the VOC detector technologies involved is also provided.

1. Introduction

Volatile Organic Compounds (VOCs) are widely used in industries as solvents or chemical intermediates. Unfortunately, they include components which, if present in the atmosphere, may represent a risk factor for human health. VOCs are also found as contaminants or as byproducts of many processes, that is, in combustion gas stacks and groundwater clean-up systems. Benzene, for example, is highly toxic beyond a Time-Weighted Average (TWA) limit of 0.5 ppm (parts per million), as compared, for instance, with the TWA limit for gasoline, which is in the range of 300 ppm. Detection of VOCs at subppm levels is, thus, of paramount importance for human safety, and, consequently, critical for industrial hygiene in hazardous environments.

The most commonly used portable field instruments for VOC detection are the hand-held Photo-Ionisation Detectors (PIDs), which may be fitted with prefilter tubes for specific gas detection. The pluses are that PIDs are accurate to subppm levels and measurements are fast, in the range of one or two minutes; for these reasons hand-held PIDs are well-suited to on-field operation. However, they have two drawbacks: they require skilled personnel and they cannot provide continuous monitoring. Wireless hand-held PIDs have recently become available on the market, thus overcoming these limitations, but they have a limited battery life, in addition to being relatively costly. This paper describes the implementation and on-field results of an end-to-end distributed monitoring system using VOC detectors which are capable of performing real-time analyses of gas concentrations in potentially hazardous sites on an unprecedented time/space scale [1].

Wireless sensor networks (WSNs), equipped with various gas sensors, have been actively used for air quality monitoring in the first decade of the 2000s [2–4]. WSNs have the advantage of offering full coverage of the terrain under

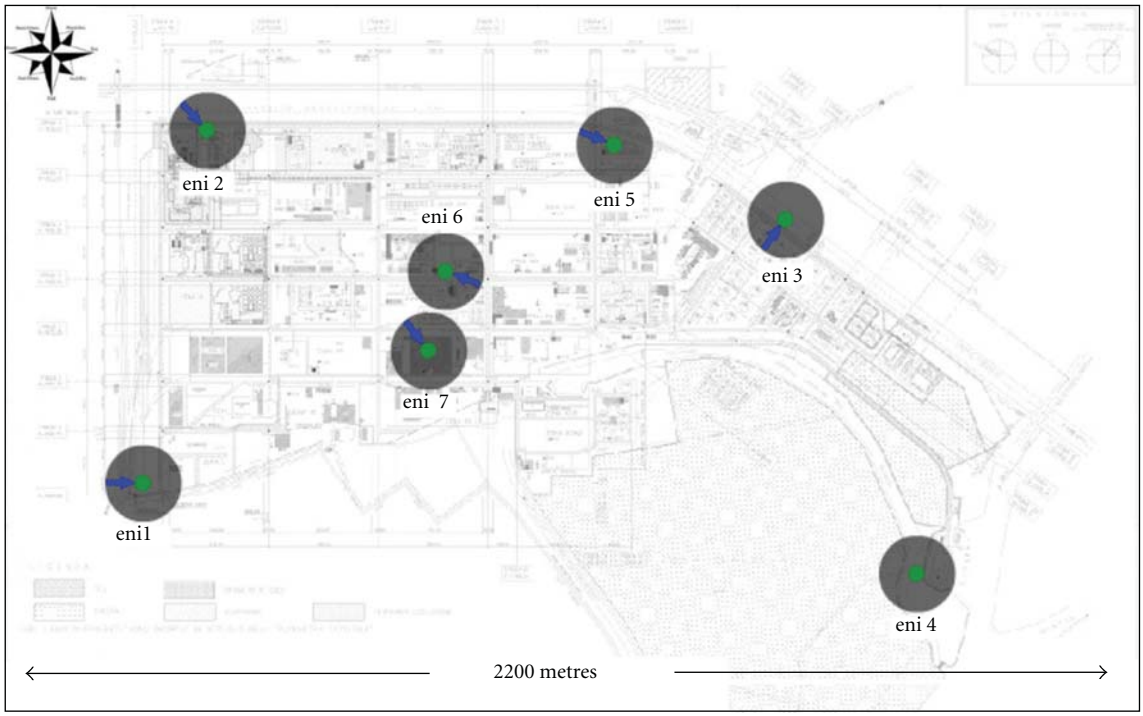


Figure 1: Installation overview. The grey circles indicate the position of each SNU; the blue arrows show wind direction.

inspection by collecting measurements from redundant portions of the zone. WSNs are thus the ideal instrument for specific and efficient environmental VOC monitoring [5, 6]. This paper describes the implementation of just such a system: a distributed network for precise VOC monitoring installed in a potentially hazardous environment. The system consists of a WSN infrastructure with nodes equipped with both weather/temperature sensors as well as VOC detectors and fitted with TCP/IP over GPRS Gateways to forward the sensors' data via Internet to a remote server. A user interface then provides access to the data as well as offering various formats of data rendering. This prototype was installed in the eni Polimeri Europa (PEM) chemical plant in Mantova, Italy, where it has been in continuous and unattended operation since April 2011. This pilot site is testing and assessing both the communications and the VOC detection technologies.

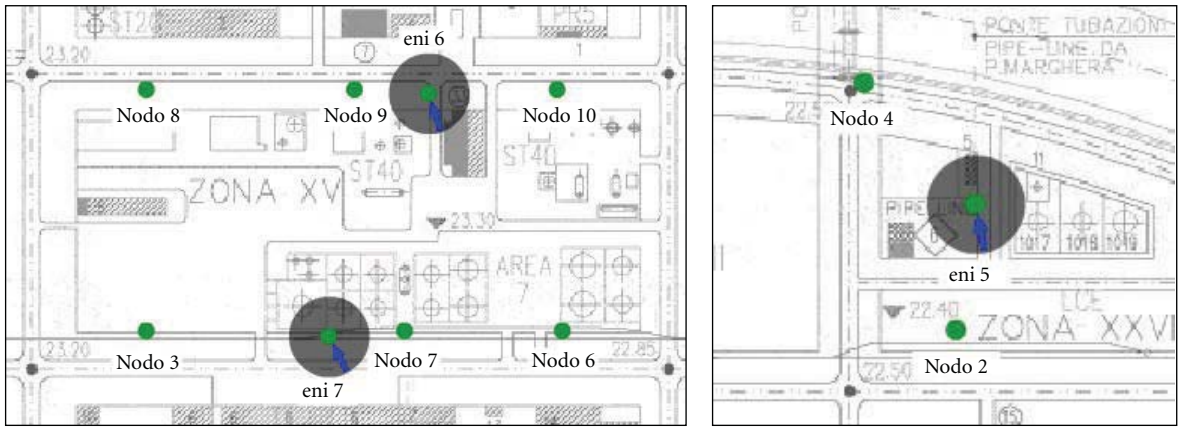
To avoid excavations, a stand-alone system, that is, one relying only on autonomous energy and connectivity resources, was designed and installed. In terms of energy requirements, the VOC detectors proved to be, by far, the greatest energy user, compared to the computational and communication units. So, to ensure a sustainable battery life for the deployed units, efficient power management strategies were studied and implemented; moreover, the WSN elements were equipped with a secondary energy source, consisting of a photovoltaic panel.

2. System Overview

A general overview of the deployed system is represented in Figure 1. First off, representative locations were identified along the perimeter of the industrial area, along with several specific internal sites where hazardous emissions might potentially occur. Owing to the extension and complexity of the Mantova plant, covering some 300 acres and featuring complex metallic infrastructures, it was decided to subdivide the area involved in the piloting into 7 different subareas. Each subarea is covered by a subnetwork consisting of a Sink Node Unit (SNU) equipped with meteorological sensors, such as wind speed/direction and relative air humidity/temperature (eni1 to eni7 in Figure 1). In addition, the eni2 unit is further equipped with a rain gauge and a solar radiation sensor.

Each SNU is connected to one or more End Node Units (ENUs) equipped with VOC detectors (see Figure 2 for an example of a configuration), appropriately distributed across the plant's property. This modular approach allows the system to be expanded and/or reconfigured according to the specific monitoring requirements, while providing redundancy in case of failure of one or more SNUs. Since the potential sources of VOC emissions in the plant are located in well-identified areas, such as the chemical plant and the benzene tanks, the deployment strategy includes a number (6) of VOC sensors surrounding the chemical plant's infrastructure, thus resulting in a virtual fence capable of effectively evaluating VOC emissions on the basis of the concentration pattern around the plant itself.

The SNUs forward meteorological data, as well as VOC concentration data, to a remote server; as noted above, Internet connectivity is provided via TCP/IP over GPRS using a GSM mobile network. Wireless connectivity uses a UHF-ISM unlicensed band. Electrical power is provided by both primary sources (batteries) and secondary sources (photovoltaic cells), as mentioned above.



(a) (b)

Figure 2: Closeups of SNU and ENU deployment around one of the chemical plants (a) and the pipeline (b). Maps are oriented according to plant's axes rather than cardinal directions.

VOC concentration and weather/climatic data are updated every minute. This intensive sampling interval allows the evolution of gas concentrations to be accurately assessed. Furthermore, when all of the weather-climatic measurements are collated, they provide a map of the area's relative air humidity/temperature (RHT) and wind speed/direction (WSD), which are crucial for providing accurate VOC-sensor read-out compensation [7]. The need for so many wind stations across the plant property is warranted by the turbulent wind distribution in the area, as can be observed by the different orientations of the blue arrows representing wind direction in Figure 1.

Three of the ENUs, eni1, eni2, and eni3, were deployed along the perimeter of the plant to locally monitor VOC concentration while correlating it with wind speed and direction; the other seven were placed around the chemical plant and in close proximity of the pipeline, which are possible sources of VOC emissions. Plans for extending the number of ENUs along the perimeter and to other chemical plants or potential sources of VOC emission are currently under consideration by PEM management.

In Figure 2 the layout of two of the subnetworks, one deployed around the chemical plant and one near the pipeline, are represented. The subnetwork around the chemical plant, Figure 2(a), consists of two SNUs, eni6 and eni7, equipped with weather sensors (air/wind), each connected with three ENUs spaced some fifty meters from each other. The subnetwork located in the pipeline area is shown in Figure 2(b); one of the two ENUs is located in close proximity of the end of the pipeline itself (*nodo 4*), while the other (*nodo 2*) is a bit further away. Sampling the VOC concentration at intervals of tens of meters allows the dispersion of VOC emissions to be evaluated; in addition, information about wind speed/direction allows the emission's source to be identified.

3. System Requirements

Given the hazardous and complex nature of the plant, developing a system that satisfied performance, reliability, and nonintrusiveness/obtrusiveness, as well as scalability/reconfigurability requirements, was very challenging. Furthermore, as continuous power is crucial, efficient power saving strategies had to be implemented to prolong battery life. The major issues taken into account were the following.

Data Grid. In the presence of multiple, scattered sources (as in the case of VOCs in industrial sites), it is important to implement a grid monitoring network, in order to have simultaneously available data over the whole area of the plant. Furthermore, correlation with meteorological input allows workers to better interpret the data and more accurately pinpoint major emission sources.

Real-Time Acquisition. Continuous real-time data permits workers to immediately detect and, thus, effectively manage emergencies that may occur within the perimeter of the plant.

Data Rate. It is important to have a high sampling rate (i.e., a sampling interval of one minute or less) to determine the details during short-term situations, so as to formulate the most appropriate corrective actions.

Scalability and Reconfigurability. Network scalability and reconfigurability are key issues, particularly at complex industrial sites. In addition to deploying fixed stations (e.g., on the perimeter of the plant), it may be useful to be able to move the monitoring stations into specific areas during specific processing phases with potentially higher VOC emissions (e.g., startup, shutdown, revamping, etc.).

Data Rendering. Depending on the purpose of the monitoring (e.g., emergency management versus air quality monitoring, etc.), it is useful to make not only real-time VOC concentration data available, but also to provide statistical indices or cumulative patterns as well. This is cost effective for those situations, for example, in which specific information for a particular compound is not required.

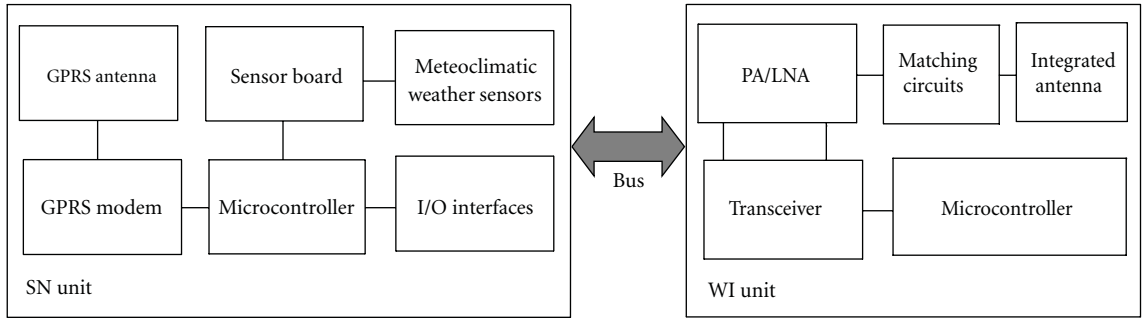


Figure 3: Block diagram of an SNU (a) and a WI unit (b).

Detection Threshold. If the purpose of monitoring is not only managing emergency situations, but also evaluating mean VOC concentrations for specific substances (in this case using the fixed monitoring stations), accordingly the choice should fall on detectors able to collect data at ppb concentration levels (as already mentioned, the air quality limit value for benzene in ambient air is currently about 1.5 ppb).

Communications. The use of wireless stations connected to a web-based graphic interface allows us to significantly reduce operating costs, as well as the costs for the infrastructure and the personnel involved.

4. The Communications Platform

The communications platform had to be able to support a scattered system of units collecting VOC emission data in real-time, while offering a high degree of flexibility and scalability, so as to allow for adding other monitoring stations as needed. Furthermore, it had to be reconfigurable, in terms of data acquisition strategies, while being more economically advantageous than traditional fixed monitoring stations.

The first requirement tackled in designing the platform was that of providing the SNUs with ubiquitous and reliable connectivity to Internet. At an early stage of the project, taking advantage of Ethernet, which was already available at some locations within the plant and which would use wireless access points to provide the SNU's with connectivity, was considered. This solution was dismissed, however, as it would have severely restricted the network's reconfigurability. The extent of the plant area, which required connectivity over several hundred square meters was another issue. Perhaps more important, however, was the presence of obstacles, such as trucks and metallic structures that could temporary or permanently affect the communication channel.

A GSM mobile network solution was decided upon as the most suitable to fulfil the above requirements; it was implemented using a proprietary TCP/IP protocol with DHCP. Dynamic reconnectivity strategies were implemented to provide efficient and reliable communication with the GSM base station. All the main communication parameters, such as IP address, IP port (server's and client's), APN, PIN code, and logic ID, can be remotely controlled. As for the wireless connectivity between the SNUs and ENUs, an unlicensed ISM UHF band (868 MHz) was selected. We shall now briefly explain the main components of the platform, the SNUs and the ENUs.

4.1. The SN and WI Units. In principle, a generic SNU consists of five components: a sensor unit, an analogue digital converter (ADC), a central processing unit (CPU), a power unit, and a communication (Wireless Interface) unit. The communication unit's task is to receive a command or a query and then transmit that data from the CPU to the outside world. The CPU is the most complex unit; it interprets the command or query to the ADC, monitors and controls power if necessary, processes received data, and manages the EN wake-up.

The block diagram of the SNU is represented in Figure 3(a). It consists of a GPRS antenna and a GPRS/EDGE quadriband modem, a sensor board, an I/O interface unit, and an ARM-9 microcontroller operating at 96 MHz.

The system is based on an embedded architecture with a high degree of integration among the different subsystems. The unit is equipped with various interfaces, including LAN/Ethernet (IEEE 802.1) with TCP/IP protocols, USB ports, and RS485/RS422 standard interfaces. The sensor board is equipped with 8 analogue inputs and 2 digital inputs. The SNU is also equipped with a Wireless Interface (WI), shown in Figure 3(b), which provides connectivity with the ENUs.

The WI, Figure 3(b), provides short-range connectivity. The WI operates on low-power, ISM UHF unlicensed band (868 MHz) with FSK modulation; it features proprietary hardware and communication protocols. Distinctive features of the unit are the integrated antenna, which is enclosed in the box for improved ruggedness, as well as a PA + LNA for a boosted link budget. The PA delivers 17 dB m to the antenna, while the receiver's noise figure was reduced to 3.5 dB, compared with the intrinsic 15 dB NF of the integrated transceiver. As a matter of fact, a connectivity range in line-of-sight in excess of 500 meters was obtained, with a reliable communication with a low BER, even in hostile EM environments.

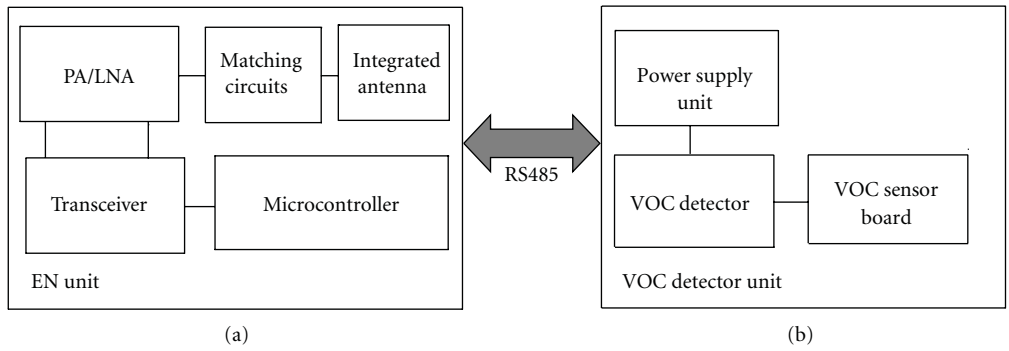


Figure 4: Block diagram of the End Node Unit (left) and the VOC detector unit (right).

The energy required for the unit's operation is provided by an 80 A h primary source and by a photovoltaic panel equipped with a smart voltage regulator. Owing to its careful low-power design, the unit could be powered with a small (20 W) photovoltaic panel while maintaining continuous, unattended operation.

4.2. The EN Unit. The block diagram of the EN is shown in Figure 4; it consists of a WI, similar to that previously described, and includes a VOC sensor board and a VOC detector. The acquisition/communication subsystem of the ENU is based on an ARM Cortex-M3 32-bit microcontroller, operating at 72 MHz, which provides the necessary computational capability on the limited power budget available.

To reduce the power requirement of the overall ENU subsystem, two different power supplies have been implemented, one for the microcontroller and one for the peripheral units; the microcontroller is able to connect/disconnect the peripheral units, thus preserving the local energy resources. The VOC detector subsystem in particular is powered by a dedicated switching voltage regulator; this provides a very stable and spike-free energy source, as required for proper operation of the VOC detector itself.

The communication between the ENU and the VOC detector board is based on an RS485 serial interface, providing high-level immunity to interference as well as bidirectional communication capability, which is needed for remote configuration/reconfiguration of the unit.

5. WSN Issues

5.1. Network Structure and Routing Schemes. Among the different alternatives, a hierarchical-based routing scheme was selected based on the particular nature of the installation: the extended area of the plant, the few critical areas of potential sources of emissions requiring a dense deployment, and the highly uneven distribution of nodes over the area. A hierarchical-based routing scheme fits the projected deployment layout well. As said before, the installation was partitioned into subnetworks to be deployed around the critical sites, with one SNU for each individual subnetwork. In principle, wireless connectivity between the SNUs could have been implemented, using one specific SNU as a gateway to the Internet. This option, however, conflicted with at least two of the major requirements. The first is the need for redundancy in case of failure of the gateway unit; in this scenario, in fact, Internet connectivity would be lost, with consequent loss of the real-time updating capability, which is considered a mandatory requirement of the system. The second need which would not have been met is that of providing full connectivity among the individual SNUs under conditions where line-of-sight propagation was not guaranteed, due to the presence of such temporary obstacles as trucks or maintenance infrastructures. A multiple GPRS gateway approach overcomes those limitations; even in the case of failure of one or more gateway units, Internet connectivity would be provided by the others still in operation, while the issue of the obstacles is circumvented. As for the wireless connectivity, a star configuration was preferred to a mesh configuration, given the limited number of nodes and the need to keep latency at a minimum.

5.2. Protocols and WSN Services. Two levels of communication protocols, in a mesh network topology, were implemented. The upper level handles communications between the SNs and the server; it uses a custom binary protocol on top of a TCP layer. This level was designed and calibrated for real-time bidirectional data exchange, where periodic signaling messages are sent from both sides. Since our sensor network necessitates a stable link, quick reconnection procedures, for whenever broken links should occur, were especially important. To ensure minimal data loss, the SNs have non-volatile data storage, as well as automatic data packet retransmission (with timestamps) after temporary downlink events. Furthermore, this design is well suited for low-power embedded platforms like ours, where limited memory and power resources are available. In fact, our protocol stack currently requires about 24 KB of flash memory (firmware) and 8 KB RAM.

The lower level, in contrast to the upper one, concerns the local data exchange between the network nodes. Here a cluster tree topology was employed; each node, which both transmits and receives data packets, is able to forward packets from the surrounding nodes when needed. In this specific application, the topology and routing schemes are based on an ID assigned to each EN unit, where the ID can be

Figure 5: Block diagram of the ENU (left) and the VOC detector unit (right) connectivity and power supply.

easily adjusted using selectors on the hardware board. This choice allows for easy support and maintenance, even when nonspecialized operators have to install, reinstall or service one or more units.

5.3. Reliability and Energy Balance. The reliability of the system and capability for stand-alone operation are summarised in Figure 5, where the main connectivity and power-supply parameters are displayed for one of the SNUs.

The energy balance and battery voltage are charted in the upper panel, from left to right, for one, three and seven days, from October 13th to 20th; hence, the net energy balance was obtained for the middle of October. GPRS connectivity performance, displayed in the middle panel, shows a reliability in excess of 99%. Wireless connectivity and data communication performance are displayed in the lower panel; no disconnection or data loss is observed in that period.

6. Energy Budget Issues

Energy budget plays a key role in the maintainability of the WSN [8]. In our case this is made even more critical by the necessity for stand-alone operation, as well as due to periodic maintenance intervals exceeding four months. Since electrical energy from the plant could not be used, secondary sources had to be locally available; photovoltaic panels (PVP) fit the bill. The SNUs are almost all equipped with PVPs, as they have to support a number of functions, including connectivity and data collection from sensors. The ENUs, when equipped with low-energy demanding sensors, have 3 to 5 years of battery life using primary sources [9]. However, in this installation the ENUs have to support the power-hungry VOC sensors. For this reason, the ENUs are also equipped with PVPs.

6.1. ENU Energy Budget. The EN nodes have been fully deployed since July 2011; since that time we have noticed that the VOC sensor energy budget is predominant compared to that of the computational/communication unit. This is a critical issue for the ENUs, as the PIDs used for reading the VOC concentration need to be continuously on to operate efficiently. This corresponds to a current draw of some 30 mA, corresponding to 720 mA h a day, more than twice the amount the communication/computational units, with their power consumption of some 360 mW a day, require. The ENU's primary source capacity is 60 Ah, which provides more than 2 full months of continuous operation.

To rely on autonomous energy resources, while providing continuous operation, a secondary energy source was integrated into the ENU in order to supply the 360 mW average required power. A 5 W photovoltaic panel can fulfil the task only under ideal sunlight conditions, that is, in summer, but hardly at all in winter. The photovoltaic power supply unit includes a charge regulator which was specifically designed to provide maximum energy transfer efficiency from the panel to the battery under any operative condition. Furthermore, the secondary energy source plays a key role in ensuring the stand-alone and unattended operation of the communication platform.

In Figure 6 the battery voltage plots are shown for the ENUs connected to SN 5 and 6. As it can be observed, the ENUs exhibit quite satisfactory charge conditions. ENU 4 (eni5 nodo 4) exhibits a slightly lower voltage level, probably due to a deployment in a partially shadowed area. The battery voltage remains above a 12.3 V value, with a slight decreasing

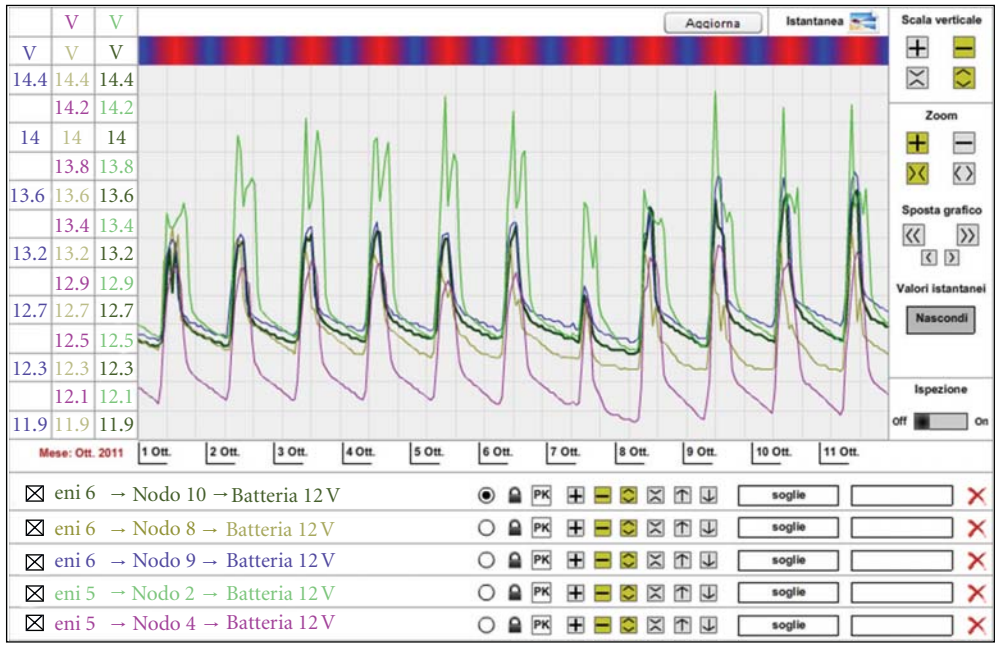


Figure 6: Battery voltage of the ENUs of the eni6 and eni5 subnetworks from October 1st to October 11th.

trend, possibly due to the lower solar energy because of the onset of autumn.

6.2. SNU Energy Budget. The SNUs were deployed at the PEM plant in the middle of April 2011; they have much higher energy requirements than the ENUs as they have to supply energy both for wireless connectivity and sensor operation.

The average current draw is around 90 mA, corresponding to a power consumption of about 1 W. SNUs have superior primary and secondary resource capabilities, with a 2-month battery life relying only on the primary source.

Figure 7 shows the battery voltage for the eni2 to eni7 SNUs; the eni1 plot is missing as the graph can only represent 6 graphs in each diagram. As can be observed, all the units demonstrate a minimum voltage exceeding 12.3 V, which denotes a satisfactory charge condition. In the above period there is also a slight decrease in the minimum battery voltage value, showing an energy imbalance between the primary and secondary sources, mostly due to sunlight reduction.

Detailed information about the charge status and trending are also available; in Figure 8 the current drawn by or supplied to the battery is compared with the charge status of the primary source. In this case, the energy balance keeps the battery voltage at a steady satisfactory level. Extensive data logs and reports are available to help the maintenance team in evaluating any critical event or service required to keep the system in full operation.

7. The VOC Detector

The VOC detector is a key element for the monitoring system's functionality. For this application two criteria were considered mandatory. The first is that the VOC detector should be operated in diffusion mode, thereby avoiding pumps or microfluidic devices which would increase the energy requirements and make the maintainability issues more critical. The second criteria was that they system should be able to operate in the very low part per billion (ppb) range, with a Minimum Detectable Level (MDL) of some 2.5 ppb with a ±5% accuracy in the 2.5 to 1000 ppb range, which represents the range of expected VOC concentration.

Table 1: System requirements.

VOC data sampling interval (minutes)	≤15
Power consumption (mW)	<200
Stabilisation time from power-on: T90 (seconds)	<60
Warm-up time (seconds)	<60
Interval between services (days)	>120
Lifetime (years)	>5
Specificity to benzene	Typically broad band

Other requirements for the VOC detectors are listed in Table 1.

An extensive analysis of the state-of-the-art devices was performed to identify the most suitable technology among the many alternatives offered by the market. Among the different candidates, which included Photo Ionisation Detectors (PID), Amperometric Sensors, Quartz Crystal Microbalance (QMC) sensors, Fully Asymmetric Ion Mobility Spectrography (FAIMS) based on MEMS, Electrochemical Sensors, and Metal Oxide Semiconductor Sensors (MOSS), it turned out that the PID technology fitted our criteria quite well, and thus it was selected as the basis for our system. The specific

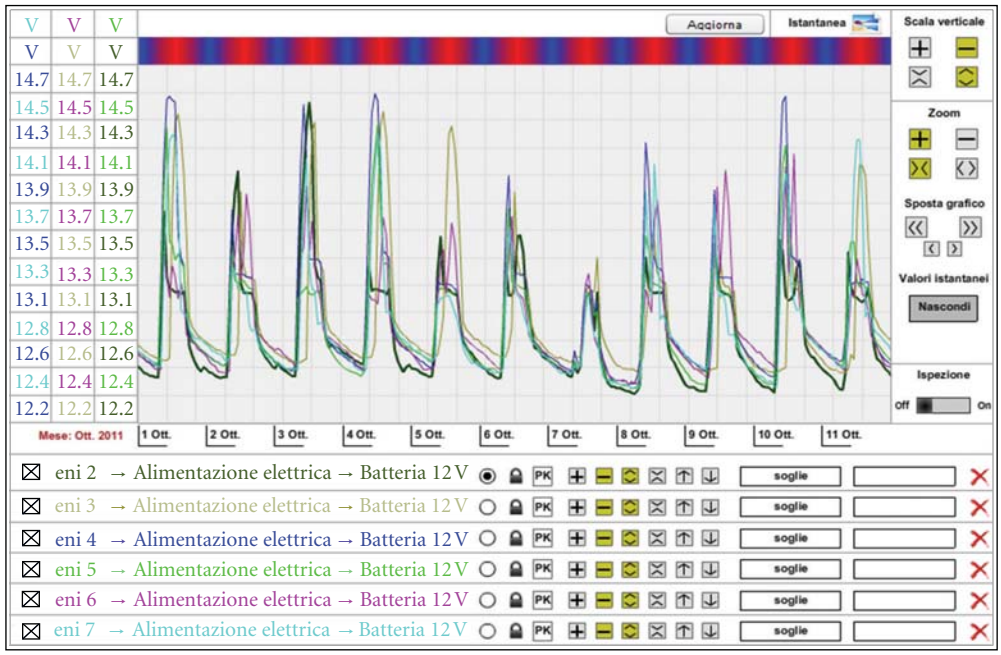


Figure 7: Battery voltage of SNUSNUs 2 to 7 in the period October 1st–October 11th.

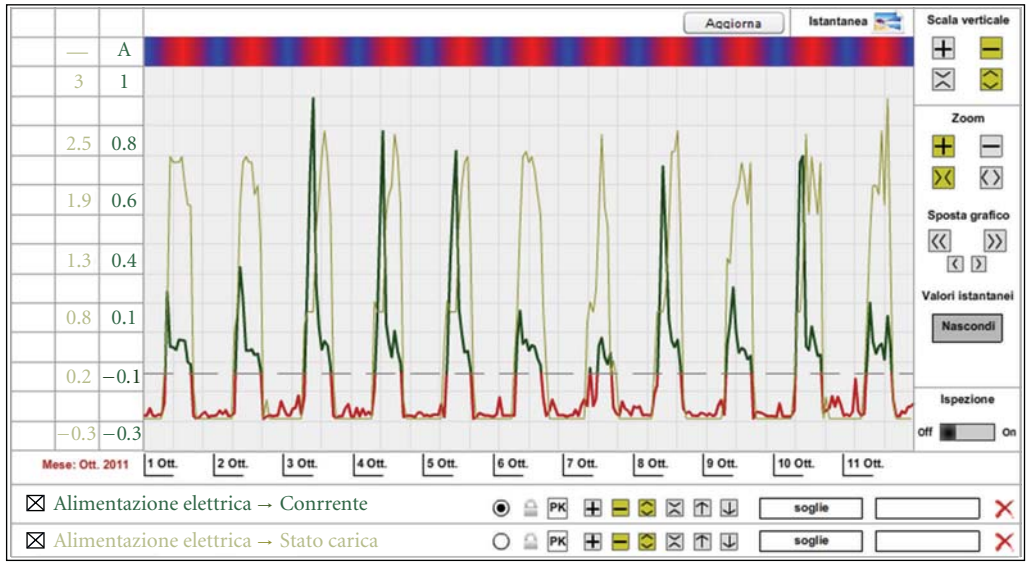


Figure 8: Current and charge status of SNU 1 from October 1st to October 11th.

device chosen was the Alphasense PID AH, which has an MDL of 5 ppb (for isobutylene) and 2.5 ppb for benzene, whose response factor is 0.5.

Both theoretical and experimental investigations of PID operation [10, 11] were carried out to assess the PID AH's performance. Two major issues were identified which could potentially affect efficient use of the PID in our system. The first was that in the low ppb range the calibration curve of the PID shows a marked nonlinearity; this would require an individualized, meticulous multipoint calibration involving higher costs and complexity. The second issue was that, when operated in diffusion mode at low ppb and after a certain time in power-off, the detector required a stabilisation time of several minutes, hence it would not be able to operate at our required one-minute intervals.

Since both of the above-mentioned limitations are intrinsically related to the PID's physical behaviour, this was carefully investigated and a behavioural model of the PID was developed to explain these phenomena. Moreover, a mathematical expression of the PID calibration curve was

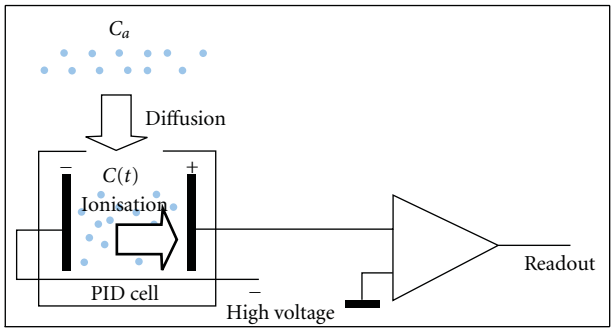


Figure 9: Basic operation of the PhotoIonisation Detector (PID).

derived. Although the theory has not been published yet [12], a brief explanation will be given here to justify the assumptions made in the design of the VOC detector unit.

With reference to Figure 9, the PID consists of a cell where an Ultraviolet (UV) lamp and two electrodes have been placed; a high voltage is established between the electrodes, generating an electric field while a hole puts the cell in direct contact with the atmosphere.

The basic operation of the PID can be described as follows. Let us assume the PID is immersed in an atmosphere containing molecules of a gas with an Ionization Potential (IP) that is lower than the energy of the photons emitted by the UV lamp; in that case an ionisation process takes place. Ions and electrons produced by the process then give rise to a drift current and are collected by the two electrodes; since each current corresponds to a specific gas concentration, the device is able to give a readout of the gas concentration within the cell. The ionisation process, by lowering the gas concentration in the cell, determines a diffusive process to take place, causing molecules to move from outside to inside the cell. It has been shown in the above-mentioned theory that at low ppb concentration levels, and when the PID is operated in diffusion mode, the rate at which the molecules diffuse is lower than that at which they are ionised. Consequently, when a stable readout is achieved, the concentration inside the cell, C_i, is somewhat lower than the concentration outside, C_a; in that condition, the readout voltage proportional to the drift current generated by the ionisation process is given by:

$$V_r = S_v C_i, \tag{1}$$

where S_v represents the PID sensitivity expressed in mV/ppm, and V_r is the read-out voltage.

The PID calibration curve defines the relationship between the measured voltage and environmental gas concentration. It is given by:

$$V(C_a) = S_v C_a. \tag{2}$$

Inspection of Formula (2) shows that, at low ppb concentrations where $C_i \neq C_a$, a correction factor is needed to appropriately map the read-out voltages onto the environmental concentration.

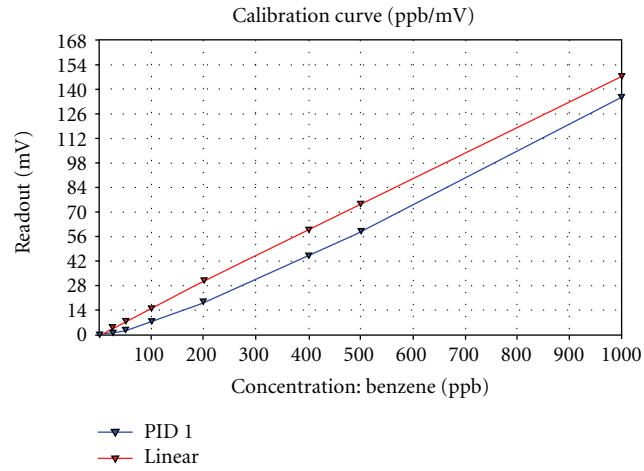


Figure 10: Experimental (blue) and linearised (red) calibration curve of the Alphasense PID AH.

7.1. PID Calibration. It is shown in [13] that at ppm concentrations the PID calibration curve is practically linear. PID calibration in the ppm concentration range is usually performed by measuring the slope of the calibration curve, S_v, at ppm concentrations; the measured read-out values are then easily mapped onto the corresponding environmental concentration values, according to (2).

This straightforward procedure cannot be used in the ppb concentration range, owing to the nonlinear behaviour of the relationship between read-out voltages and environmental concentration, thus requiringan individualized, careful and multipoint calibration to be implemented, with additional costs and complexity. To overcome this limitation, a linearisation procedure for the ppb concentration range was developed, based on a behavioural model of the PID.

Accordingly, the voltage readouts measured by the detector, $V_r(C_a)$, are first multiplied by a nonlinearity compensation factor, $\gamma(C_a)$, a function of the environmental concentration C_a, according to the following expression:

$$V(C_a) = \gamma(C_a) V_r(C_a). \tag{3}$$

As a result, a linear calibration curve is obtained for the ppb concentration range, allowing the environmental concentration values to be attained by the measured read-out values, according to expression (3). The result of this process is shown in Figure 10, where the experimental function $V_r(C_a)$ is represented by the blue line, and the results of the linearisation process are represented by the red line.

7.2. PID Stabilisation Time. The second problem arising from PID operation is represented by the stabilisation time required to achieve a stable read-out time in diffusion mode at low concentrations (tens or hundreds of ppb), which represents the area of operation of the VOC detectors for our application. This effect is clearly shown in the experiment in Figure 11.

In the experiment, C_a was around 50 ppb, which represents the average concentration where the PID is supposed

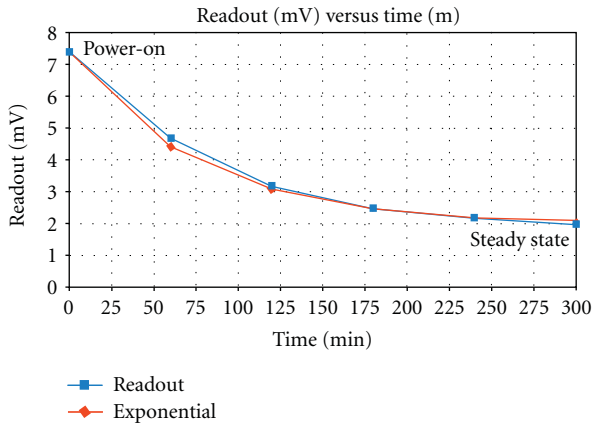


Figure 11: PID stabilisation time after power-on.

to be set up. At time $t = 0$, the PID was powered-on and the readout was sampled at a 1 minute interval. The PID readout (blue line), is compared with an 80-second time-constant exponential curve (red line), showing a very good agreement. The reason for the relatively high stabilisation time at low-ppbs is related to the same phenomenon that gives rise to a nonlinear calibration curve, that is, that a stable readout is obtained when the concentration inside the cell is lower than the environmental concentration. This effect is clearly observed in Figure 10; at power-on the read-out voltage of the ionisation process is proportional to C_a, which is higher than C_i. With the progression of the ionisation process, C_i decreases and eventually reaches a stable value.

The PID's stabilisation time highly impacts the device's operation. According to the specifications of Table 1, a minimum VOC data sampling interval of at least fifteen minutes is required. Furthermore, a duty-cycled operation is desirable, in principle, to prolong the battery life. Operating the detector on a fifteen minute duty-cycle basis, and taking into account the stabilisation time, the benefit in terms of maintenance effort is marginal and has to be compared with the advantage of achieving a more time-intensive monitoring of VOC concentration, as provided by continuous power-on operation. Accordingly, it was decided to operate the VOC detectors continuously at one-minute data sampling; this decision proved to be very effective, as some emission events at the plant show very rapid variation, which could result in a difficult interpretation at ten-minute sampling rates.

Continuous power-on operation requires a 35 mA h charge, which corresponds to about 2 months of full operation with a 60 A h primary energy source. On the other hand, the UV lamp's expected life is more than 6000 hours of continuous operation, that is, four months at least. Consequently, a four-month routine maintenance is planned for the system, so that UV lamp replacement, PID refitting, and battery replacement can be appropriately scheduled.

8. Experimental Results

Data gathered from the field are forwarded to a central database for data storage and data rendering. A rich and proactive user interface was implemented in order to provide detailed graphical data analysis and presentation of the relevant parameters, both in graphical and bidimensional format. Data from the individual sensors deployed on the field, either microclimatic or VOC, can be directly accessed and presented in various formats.

Information at a glance is provided by the map in Figure 12(a). As explained in Figure 1, the SNUs of the individual subnetworks are represented as grey circles with a blue arrow indicating the wind direction. By positioning the mouse pointer over an SNU icon, all of its microclimatic and VOC parameter values are displayed. As shown in Figure 2, a closeup of subnetworks can be accessed, for detailed information of VOC concentration readout at each ENU.

A summary of the sensor status for each deployed unit can be obtained by opening the summary panel, Figure 12(b). The summary panel reports current air temperature/humidity values, along with min/max values of the day (lower left in Figure 12(b)), wind speed and direction (upper left in Figure 12(b)), and VOC concentration (bar chart on right-hand side of Figure 12(b)) over the last six hours.

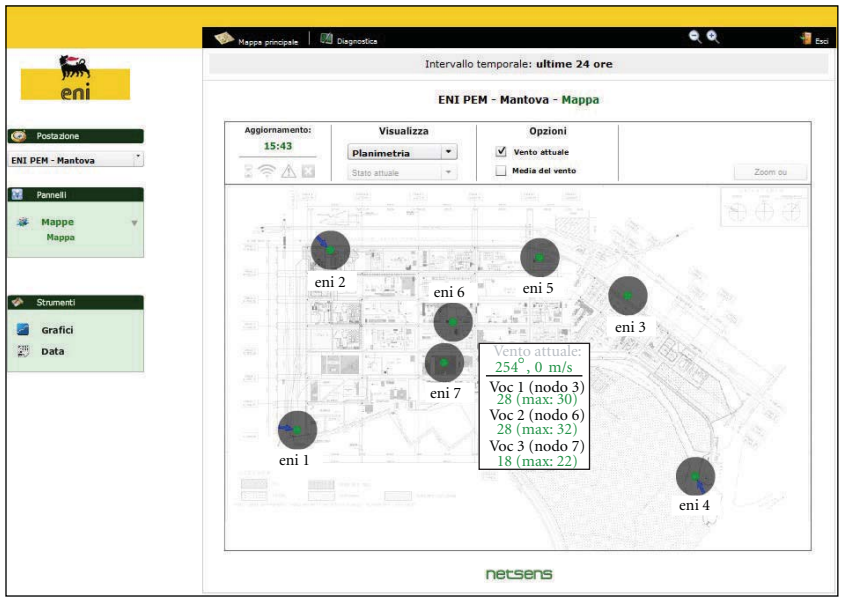
In Figure 13 a comprehensive 2D representation of the VOC, Figure 13(a), and climatic, Figure 13(b), parameter distribution over the plant is shown.

The representations were obtained using an interpolation algorithm in pseudocolour. Blue denotes a lower concentration/temperature, while red indicates a higher one. It should be emphasized that the choice of red was merely a chromatic one; it has absolutely no reference to any risky or critical condition.

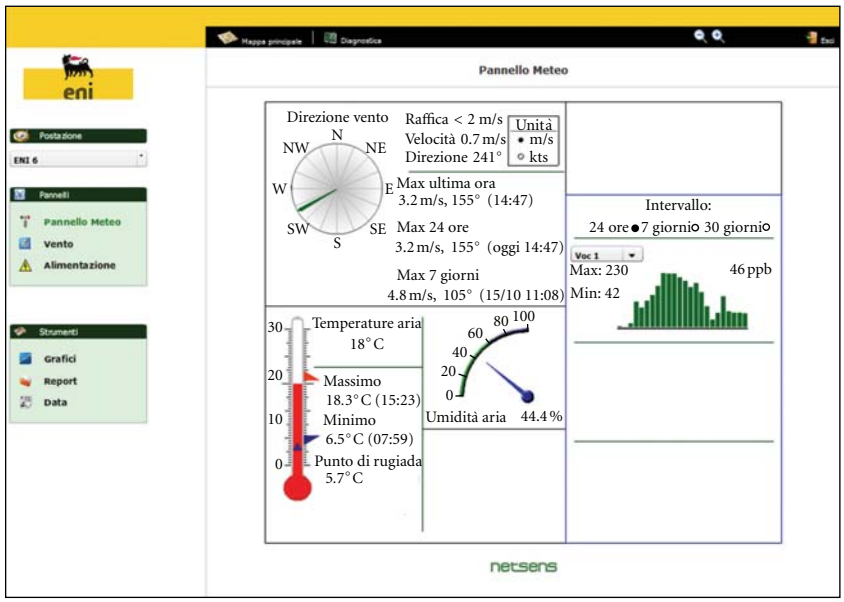
A graphic representation of data gathered by each sensor on the field can be obtained by opening the graphic panel window, see Figure 14. This panel allows anyone to display the stored data in any time interval in graphic format; up to six different and arbitrarily selected sensors can be represented in the same graphic window for purpose of analysis and comparison. In Figure 14, for example, the six traces represent the VOC concentration values detected by the six PIDs deployed around the chemical plant. The optimum uniformity among the background concentration levels should be noted, as it demonstrates the effectiveness of the calibration procedure.

The background values range from 50 to 150 ppb. Thanks to the intensive 1-minute sample interval, the evolution of the concentration, along with other relevant weather parameters, can be accurately displayed. It should be noted that the spikes that can be observed in the traces have an average duration of 2-3 minutes.

When VOC sources need to be identified, the correlation between wind/speed direction and VOC concentration is vital. In Figure 15 two different representations of VOC concentrations combined with the wind direction data are shown for two detectors located at the plant perimeter, namely, ENUs 1 and 2. The plot, in polar coordinates, represents the wind directions referenced to the North and the VOC concentration in ppb. The diagram shows the average VOC concentration in all directions, as detected by



(a)



(b)

Figure 12: Examples of data rendering: installation map (a) and summary panel (b).

the PID during a full day. The plot gives an overview of the predominant orientation of the VOC flux during the day.

In Figure 15(a) the diagram of ENU 1, located in the southwest corner of the plant, is shown. The area outside the plant is to the west, while the area inside is to the east of ENU 1. In Figure 15(a) the VOC concentration is higher in quadrants I and IV, showing that the net VOC flow is entering the plant area. This may be related to the emissions generated by the traffic on the motorway running along the west side of the plant, or possible emissions from other industrial sites.

Across the motorway there are also a petroleum refinery (WSW) and its storage area (WNW). The same diagram for ENU 2, which is instead located at the northwest corner of the plant, is shown in Figure 15(b). Here the opposite situation is found, as the highest concentration values are in the II quadrant. It should be noted, however, that the data for ENU 2 was recorded on a Sunday when the traffic is assumed to be less intense.

9. Conclusions

An end-to-end distributed monitoring system of integrated VOC detectors, capable of performing real-time analysis of gas concentration in hazardous sites at an unprecedented

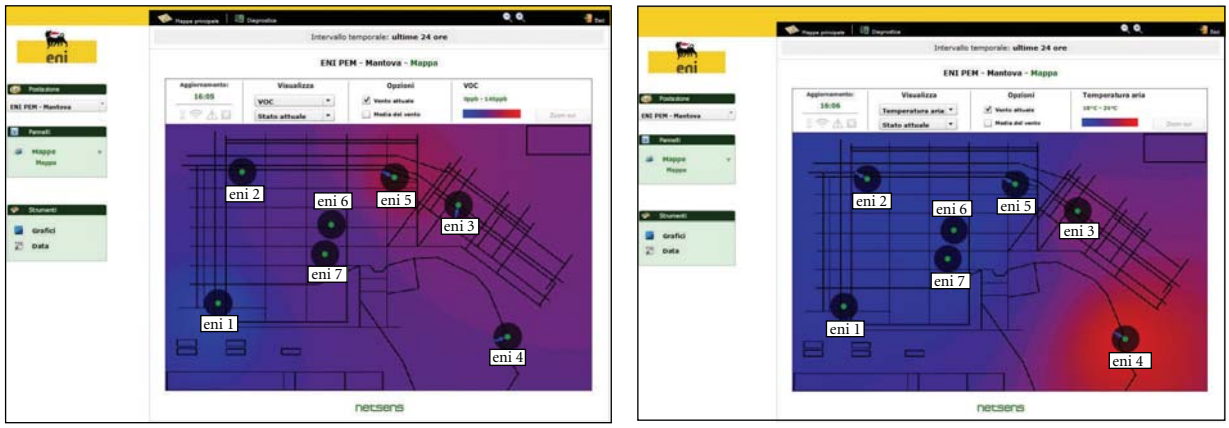


Figure 13: Examples of data rendering: VOC concentration (a) and temperature distribution (b) across plant site.

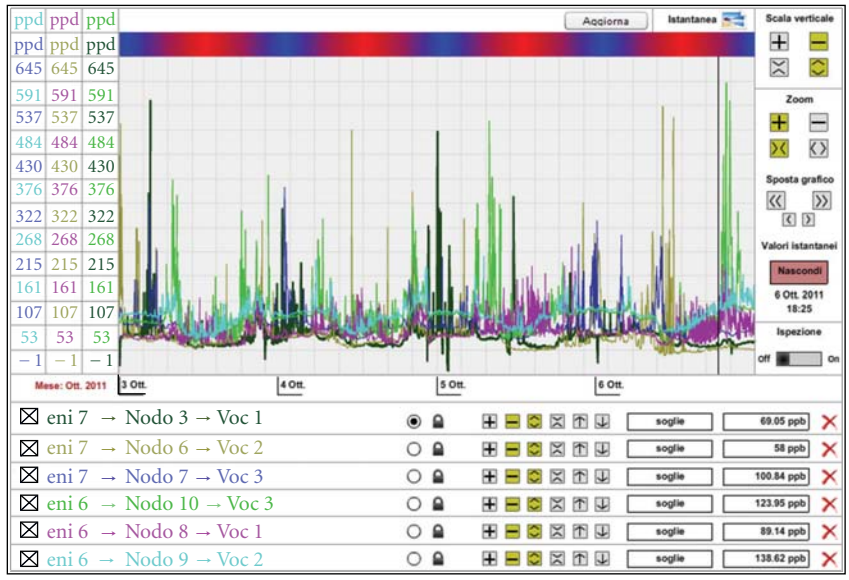


Figure 14: Example of data rendering: Graph of VOC concentration in the six detectors deployed around the chemical plant.

time/space scale, has been implemented and successfully tested in an industrial site. The aim was to provide the industrial site with a flexible and cost-effective monitoring tool in order to achieve a better management of emergency situations, to identify emission sources in real time, and to collect continuous VOC concentration data using easily redeployable and rationally distributed monitoring stations.

The piloting of the system allowed us to pinpoint key traits. Collecting data at 1-minute time intervals meets several needs: identifying short-term critical events, quantifying the emission impacts as a function of weather conditions as well as of operational process, in addition to identifying potentially VOC sources in the plant area. Moreover, the choice of a WSN communication platform gave excellent results, above all in allowing for redeploying and rescaling the network's configuration according to specific needs as they arise, while, at the same time, greatly reducing installation costs. Furthermore, real-time data through a web-based interface allowed both adequate levels of control and quick data interpretation in order to manage specific situations. In terms of the actual detectors, among the various alternatives available on the market, PID technology proved to meet all the major requirements, as PIDs are effective in terms of energy consumption, measuring range, cost, and maintenance once installed in the field. Finally, fitting weather sensors at the nodes of the main network stations allowed for a clearer understanding of on-field phenomena and their evolution, thus providing accurate identification of potential emission sources.

Future activity will include a number of further developments, primarily the development of a standard application to allow the deployment of WSN in other network industries

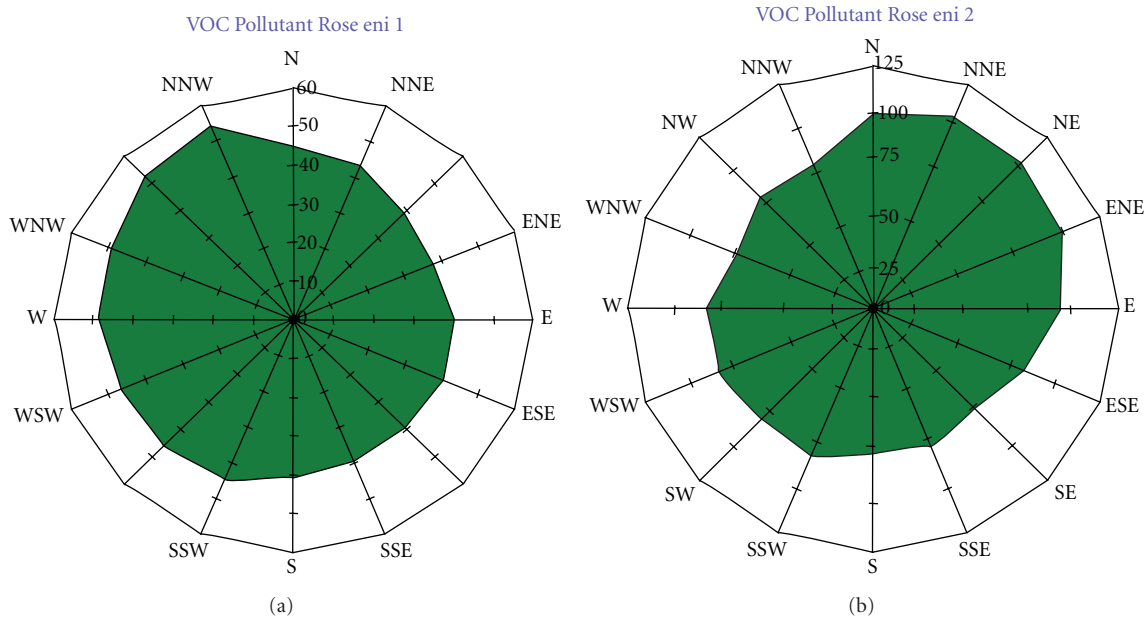


Figure 15: Example of data rendering: correlation between wind and VOC concentration for ENU 1 (a) and 2 (b).

(e.g., refineries) in addition to an assessment of potential applications for WSN infrastructure monitoring of other environmental indicators.

Acknowledgments

This work was supported by eni S.p.A. under contract no. 3500007596. The authors wish to thank W. O. Ho and A. Burnley, at Alphasense Ltd., for their many helpful comments and clarifications concerning PID operation, as well as E. Benvenuti, at Netsens s.r.l., for his valuable technical support. Assistance and support from the management and technical staff of Polimeri Europa Mantova is gratefully acknowledged. The valuable support by Mrs J. A. Thonn for revising the English is also acknowledged.

References

[1] G. Manes, R. Fusco, L. Gelpi, A. Manes, D. Di Palma, and G. Collodi, *Real-Time Monitoring of Volatile Organic Compounds in Hazardous Sites*, chapter 14, Intech Book, Environmental Monitoring, 2011.

[2] W. Tsujita, H. Ishida, and T. Moriizumi, "Dynamic gas sensor network for air pollution monitoring and its auto-calibration," in *Proceedings of the IEEE Sensors*, vol. 1, pp. 56–59, October 2004.

[3] F. Tsow, E. Forzani, A. Rai et al., "A wearable and wireless sensor system for real-time monitoring of toxic environmental volatile organic compounds in ," *IEEE Sensors Journal*, vol. 9, no. 12, pp. 1734–1740, 2009.

[4] S. Choi, N. Kim, H. Cha, and R. Ha, "Micro sensor node for air pollutant monitoring: hardware and software issues," *Sensors*, vol. 9, no. 10, pp. 7970–7987, 2009.

[5] R. Szewczyk, A. Mainwaring, J. Polastre, J. Anderson, and D. Culler, "An analysis of a large scale habitat monitoring application," in *Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems*, pp. 214–226, November 2004.

[6] R. Adler, P. Buonadonna, J. Chabra et al., *Design and Deployment of Industrial Sensor Networks: Experiences from the North Sea and a Semiconductor Plant in ACM SenSys*, San Diego, Calif, USA, 2005.

[7] MiniPID User Manual V1.8, IonScience Ltd, 2000.

[8] J. Jeong, D. E. Culler, and J. H. Oh, "Empirical analysis of transmission power control algorithms for wireless sensor networks," in *Proceedings of the 4th International Conference on Networked Sensing Systems (INSS '07)*, pp. 27–34, IEEE Press, June 2007.

[9] G. Manes, R. Fantacci, F. Chiti et al., "Energy efficient MAC protocols for wireless sensor networks endowed with directive antennas: a cross-layer solution," in *Proceedings of the IEEE Radio and Wireless Conference*, pp. 239–242, Orlando, Fla, USA, January 2009.

[10] J. G. W. Price, D. C. Fenimore, P. G. Simmonds, and A. Zlatkis, "Design and operation of a photoionization detector for gas chromatography," *Analytical Chemistry*, vol. 40, no. 3, pp. 541–547, 1968.

[11] D. C. Locke and C. E. Meloan, "Study of the photoionization detector for gas chromatography," *Analytical Chemistry*, vol. 37, no. 3, pp. 389–395, 1965.

[12] G. F. Manes, unpublished results.

[13] Alphasense Ltd., Technical specifications; Doc. Ref. PID-AH/MAR11.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 975147, 14 pages
doi:10.1155/2012/975147

Research Article

Anchor-Node-Based Distributed Localization with Error Correction in Wireless Sensor Networks

Taeyoung Kim,[1] Minhan Shon,[1] Mihui Kim,[2] Dongsoo S. Kim,[3] and Hyunseung Choo[1]

[1] *School of Information and Communication Engineering, Sungkyunkwan University, 300 Cheoncheon-dong, Jangan-gu, Suwon-si, Gyeonggi-do 440-746, Republic of Korea*
[2] *Department of Computer Engineering, Hankyong National University, 327 Chungang-no, Anseong-si, Gyeonggi-do 456-749, Republic of Korea*
[3] *Department of Electrical and Computer Engineering, Indiana University-Purdue University Indianapolis, 723 W. Michigan Street SL160, Indianapolis, IN 46202, USA*

Correspondence should be addressed to Hyunseung Choo, choo@ece.skku.ac.kr

Received 2 November 2011; Accepted 18 January 2012

Academic Editor: Tai Hoon Kim

This paper proposes a scheme to enhance localization in terms of accuracy and transmission overhead in wireless sensor networks. This scheme starts from a basic anchor-node-based distributed localization (ADL) using grid scan with the information of anchor nodes within two-hop distance. Even though the localization accuracy of ADL is higher than that of previous schemes (e.g., DRLS), estimation error can be propagated when the ratio of anchor nodes is low. Thus, after each normal node estimates the initial position with ADL, it checks whether the position needs to be corrected because of the insufficient anchors within two-hop distance, that is, the node is in sparse anchor area. If correction needs, the initial position is repositioned using hop progress by the information of anchor nodes located several hops away so that error propagation is reduced (REP); the hop progress is an estimated hop distance using probability based on the density of sensor nodes. Results via in-depth simulation show that ADL has about 12% higher localization accuracy and about 10% lower message transmission cost than DRLS. In addition, the localization accuracy of ADL with REP is about 30% higher than that of DRLS, even though message transmission cost is increased.

1. Introduction

Many applications of wireless sensor networks such as object tracking, environment and habitat monitoring, intrusion detection, and geographic routing rely on the location information of sensor nodes [1]. To provide accurate position information of sensor nodes, each sensor node can be equipped with GPS [2]. In many applications, however, it is not possible that all sensor nodes are equipped with GPS economically and technically. For examples, GPS signal hardly reaches to all nodes when wireless sensor networks are deployed in underground structure or underwater environments, where a sensor node must estimate its position from various information of other nodes that are possibly equipped with GPS or estimate their positions from others as well. A localization refers to a process for determining the position of users or wireless devices without using positioning hardware device as GPS [3, 4].

Localization schemes are categorized into two classes: range-based positioning scheme and range-free positioning scheme. A range-based positioning scheme measures the distance or the angle between sensor nodes and estimate positions using triangulation, trilateration, and multilateration algorithms. Typical range-based positioning schemes include Time of Arrival (ToA) [5], Time Difference of Arrival (TDoA) [6], Angle of Arrival (AoA) [7], Received Signal Strength Indicator (RSSI) [8], and RSS-based localization using direction calibration [9]. The range-based positioning schemes require additional devices to measure the distance or the angle between sensor nodes. Some range-based positioning schemes are sensitive to temporary interference of signal such as noise and fading. They are unsuited to massive wireless sensor networks. A range-free scheme estimates positions using connectivity information between sensor nodes [10–20]. Among the range-free positioning schemes are Centroid [10], Convex Position Estimation (CPE) [11], and

Distance Vector-Hop (DV-Hop) [12] methods. The range-free positioning schemes, however, suffer from high message transmission cost to gather the connectivity information and high computational cost to estimate positions of sensor nodes. In addition, they have shown high position estimation error compared to other methods.

Distributed Range-free Localization Scheme (DRLS) [21] improved localization accuracy of normal nodes by utilizing anchor nodes within two-hop distance. The scheme used grid scanning algorithm to estimate the initial position of a normal node and tried to improve the estimation accuracy by vector-based refinement. The vector-based refinement, however, could inaccurately decide the position of normal nodes depending on the relative position of anchor nodes and needs square-root calculations requiring high computation power. In addition, a normal node can infer an incorrect localization when it estimates its position from other normal nodes due to the lack of neighbor anchor nodes. The incorrect localization of a normal node deteriorates the localization of another normal node so that the error is propagated over the networks.

This paper proposes an Anchor-based Distributed Localization (ADL) scheme and its Reducing Error Propagation (REP) scheme. The ADL devises a grid scanning algorithm of two-hop anchor nodes instead of the vector-based refinement of DRLS, which reduces the computational cost occurred by the refinement of DRLS and improves the accuracy of the localization. The ADL with REP enhances the initial estimation of a normal node that does not have anchor nodes within two-hop and reduces the impact of the error propagation frequently occurred with a sparse anchor nodes deployment. Our contributions in this paper are as follows.

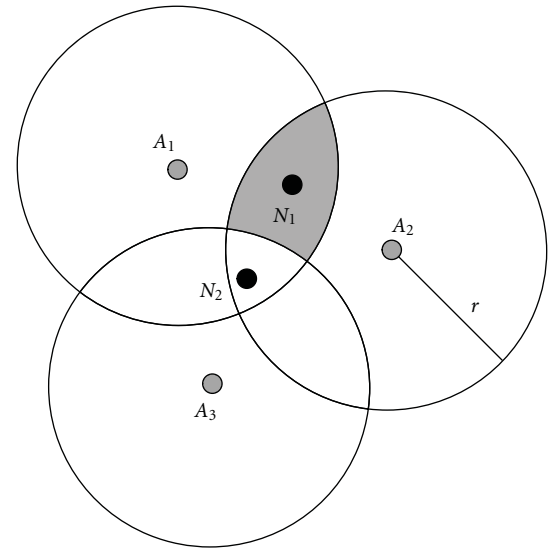
(i) We designed a basic ADL to improve the accuracy of the localization and reduce the computational cost, and proposed an REP scheme based on ADL to decrease the propagation error in the sparse anchor nodes deployment, as previous work [22, 23]. In this paper, we organize the ADL with REP and then develop the various simulation environments to evaluate our scheme from the several points of view.

(ii) The results via in-depth simulation show that the basic ADL has about 12% higher localization accuracy and about 10% lower message transmission cost than DRLS. In addition, the localization accuracy of ADL with REP is about 30% higher than that of DRLS, even though message transmission cost is increased.

The remainder of the paper is organized as follows: Section 2 presents preliminary study and previous work. Section 3 details our proposed scheme. Section 4 discusses simulation results. Finally, Section 5 concludes the paper.

2. Preliminaries and Related Work

2.1. Assumptions and Definition. The purpose of the proposed scheme is to provide more accurate localization by applying a range-free localization scheme. In the range-free localization scheme, we assume that a sensor network consists of normal nodes and anchor nodes, where the normal nodes do not know their position but the anchor nodes are able to acquire their positions via external positioning device such as GPS. All sensor nodes including the normal node and the anchor node are randomly deployed in a sensor field, and they do not move after the initial deployment. In addition, each sensor node has a unique ID and the same data transmission radius [21].



A_i: Anchor node
N_i: Normal node
r: Transmission radius

Figure 1: An example of one- and two-hop anchor nodes (A_1 and A_2 are one-hop anchor nodes of N_1 and N_2, and A_3 is a two-hop anchor node of N_1.).

In the proposed scheme, each normal node estimates its position by obtaining position information from anchor nodes within two-hop distance. An one-hop anchor node is an anchor node located within the data transmission radius r of a normal node. A two-hop anchor node is reachable by a normal node with a two-hop path alone. For an anchor node to be a two-hop anchor, the anchor is located within twice transmission range ($2r$) of the normal node, but out of a transmission range. Figure 1 shows an example of one-hop and two-hop anchor nodes. Normal node N_1 considers anchor nodes A_1 and A_2 are one-hop anchor nodes because they are located within transmission radius of N_1. Anchor node A_3 is a two-hop anchor node of N_1 that indirectly obtains the position information of A_3 via a normal node N_2.

2.2. Related Work. In the range-free localization scheme, a normal node estimates its position using connectivity information between itself and its neighbor nodes. Thus, the scheme requires data transmission instead of additional devices to measure distances or angles between sensor nodes. In Centroid [10], each normal node estimates its position by calculating the average position of the neighbor anchor nodes, after gathering the position information of neighbor

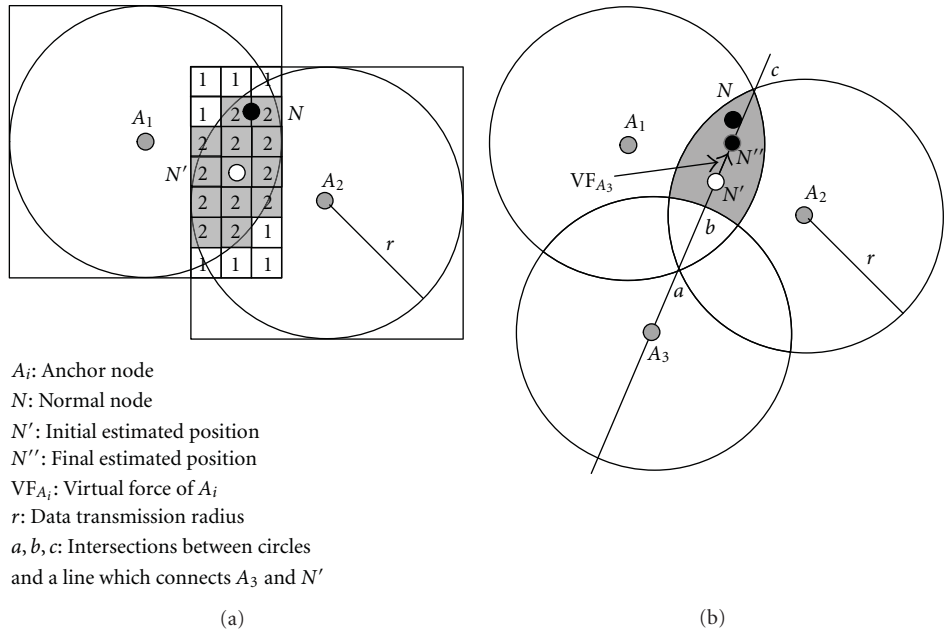


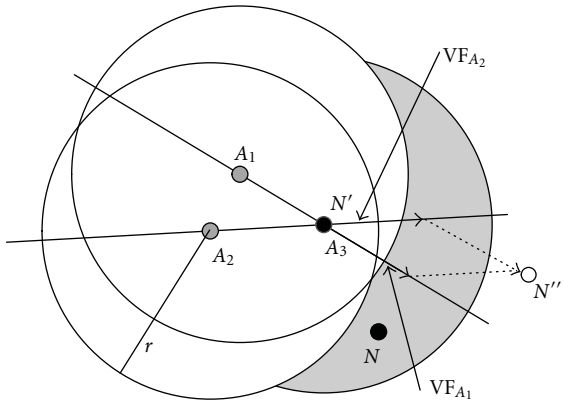
Figure 2: An example of the position estimation by DRLS: (a) initial position estimation by grid-scan; (b) refinement of initial position.

anchor nodes. This localization method is very simple but its accuracy becomes very low when there are a few sensor nodes in the field. In addition, its localization accuracy is highly depending on the deploying pattern of anchor nodes.

Convex position estimation (CPE) [11] proposed to relieve the wrong localization issue of the Centroid based on the position of anchor nodes. CPE also uses the connectivity information between sensor nodes, where each normal node assumes an imaginary rectangle, or an estimative rectangle (ER) that is tangent to an overlapped region of transmission radiuses of neighbor anchor nodes. The normal node should be on ER and the center of the rectangle is approximated as the estimated position of the normal node.

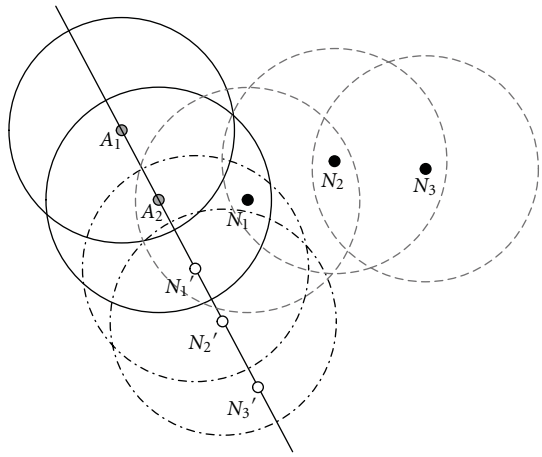
Distributed Range-free Localization Scheme (DRLS) [21] extends to use anchor nodes within two-hop distance. The scheme consists of two phases as *initial placement* (i.e., *grid scan*) and *refinement*. In the initial placement phase, the transmission range of each anchor node is virtually considered as a square instead of a circle and a normal node N assumes an ER as a region overlapped by the squares of anchor nodes within one-hop distance from the normal node. The ER calculated by this method covers a larger area than the one of CPE does. To compensate this looseness, the scheme applies a grid scanning so that the ER is divided to small grid-shaped cells. The number of times each cell overlaps with the transmission range of anchor nodes is calculated. The gray cells in Figure 2(a) have the highest value that is the number of nesting and the normal node N should be on these cells. The average position of gray cells, N', is estimated as the initial position. If a normal node is in two-hop path of an anchor node, the second phase allows refining its position. Figure 2(b) shows the refinement phase of DRLS. Because the normal node N cannot be in the transmission radius of anchor node A_3 that is within two-hop distance from N, the initial estimated position N' is adjusted using a virtual force ($\mathrm{VF}_{\mathrm{A}_3}$) that is made by the effect of A_3 to the region formed by A_1 and A_2. The direction of $\mathrm{VF}_{\mathrm{A}_3}$ is from A_3 to N', and the length of $\mathrm{VF}_{\mathrm{A}_3}$ is $|\overrightarrow{VF_{\mathrm{A}_3}}| = \overline{ab} \cdot \overline{N^1c}/\overline{ac}$

However, DRLS has some problems. First of all, two-hop flooding of DRLS is based on anchor nodes. Each anchor node exchanges its position information with anchor nodes that are located within two-hop distance and provides all of information to its neighbor normal nodes. Even though normal nodes obtain some position information during the information exchange between anchor nodes, the normal nodes do not use the information. Secondly, localization accuracy can be worse depending on the deployment pattern of sensor nodes in the refinement phase as illustrated in Figure 3. N' is obtained as the estimated position of a normal node N by an one-hop anchor node A_3. Two-hop anchor nodes A_1 and A_2 decide the VF as the sum of the $\mathrm{VF}_{\mathrm{A}_1}$ and $\mathrm{VF}_{\mathrm{A}_2}$, so that N'' can be forced to be at the position where the normal node cannot be located. At the last but the most importantly, normal nodes with no neighbor anchor nodes estimate their positions with the position information of neighbor normal nodes that already have some localization errors. The error propagation causes a serious problem in many situations. Figure 4 shows an example of frequently occurring error propagation. When two anchor nodes are located together, such as A_1 and A_2, the estimated position, such as N_1', N_2', and N_3', is on the line connecting two anchor nodes, even though the real positions of normal nodes are N_1, N_2, and N_3. The error propagation is amplified by the distance from anchor nodes. In this paper, we propose a novel scheme substitutes the vector-based refinement and reduces error propagation to resolve the problems discussed above.



A_i: Anchor node
N: Normal node
N': Initial estimated position
N'': Final estimated position
VF$_{A_i}$: Virtual force of A_i
r: Data transmission radius

Figure 3: An example that localization error increases in the refinement of DRLS.



A_i: Anchor node
N_i: Normal node
N_i': Estimated position

Figure 4: An example of error propagation in DRLS when two anchor nodes are distributed together.

3. Anchor-Node-Based Distributed Localization with Error Correction

This section presents the basic ADL and advances the ADL with error correction to reduce the error propagation in sparse anchor node environments. Figure 5 outlines our proposed scheme, the ADL with ERP. Each normal node obtains an initial estimated position after performing ADL, and then it check whether the position needs to be corrected because of the insufficient anchors within two-hop, that is, the node is in sparse anchor area. If yes, each node estimates the final position by hop progress, otherwise, the initial position becomes a final value.

3.1. Basic ADL. The Basic ADL consists of three steps: *anchor node selection*, *ER construction*, and *distributed grid scan*. We introduce the steps and the limitation of ADL in sparse anchor node environments.

3.1.1. Anchor Node Selection. In the anchor node selection step, two-hop flooding is used for normal nodes to obtain the information of anchor nodes within two-hop distance. Two-hop flooding is started from each anchor node, and each anchor node broadcasts its ID and position information to its neighbor sensor nodes. All normal nodes that receive the information broadcasted decide the anchor nodes of the information obtained as their one-hop anchor nodes. Then, the normal nodes that have the information of one-hop anchor nodes broadcast the information of their one-hop anchor nodes to its neighbor normal nodes. The normal nodes that receive the information of the one-hop anchor nodes from their neighbor sensor nodes consider the anchor nodes of the information obtained as two-hop anchor nodes.

3.1.2. ER Construction. A normal node can be located in a rectangle with thick lines of Figure 6(a). However, it is very difficult to calculate this region, since the region is presented by the subsumption relationship between circles. Therefore, an approximate region termed ER is determined by rectangles that are tangential to the transmission radius of each anchor node, to find this region with lower computational cost. Rectangles, that are tangential to the circles whose centers are the positions of one-hop anchor nodes and radiuses are the transmission radius of sensor nodes, are formed to create ER. In addition, rectangles that are tangential to the circles whose centers are the positions of two-hop anchor nodes and radiuses are the two times of transmission radius, are made. Subsequently, the overlapped region of all rectangles is selected as ER. Figure 6(a) shows an example to form an ER as the overlapped region of three rectangles.

3.1.3. Distributed Grid Scan. The ER calculated by a normal node includes a region that is larger than the region in which the normal node can actually exist. Accordingly, the localization accuracy can be enhanced by excluding the region where a normal node cannot actually exist from the ER using the grid scan algorithm. After dividing the ER into a set of grid cells, the normal node scans each cell to decide if the cell region should be excluded as shown in Figure 6(b). When the data transmission radius is r, the length of the edge on each cell is $0.1r$ by considering the scanning computational cost and localization accuracy. Each cell has a binary value as a representative value initialed by 1. The binary value 1 denotes that the normal node can exist in the cell. While scanning the cells, the normal node excludes cells in which the normal node cannot be located (i.e., setting 0 to the excluded cells). The center of gravity on each cell is deemed to be the representative position of the cell. Finally, normal node N obtains final estimated position N' by computing

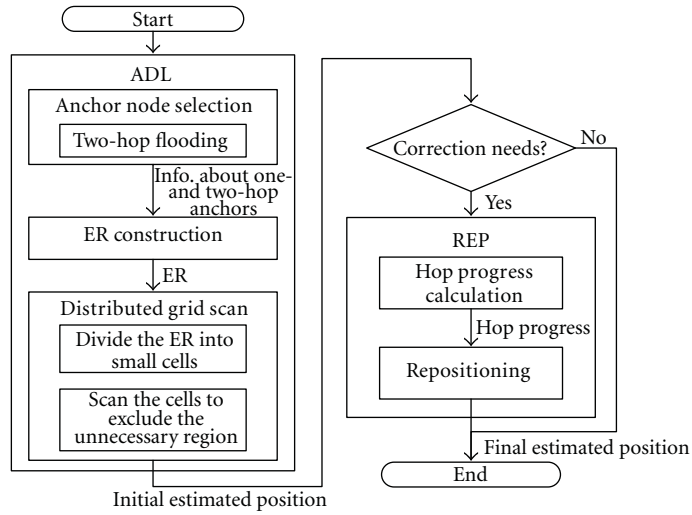


Figure 5: Outline of proposed localization scheme.

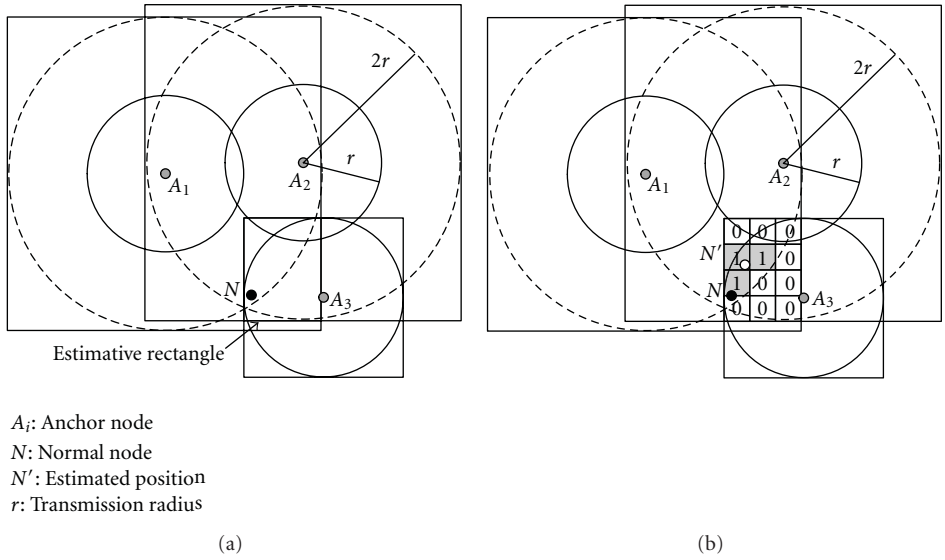


Figure 6: An example of ER construction and grid scan: (a) ER construction; (b) grid Scan.

the average of the representative coordinate values in each shaded cell with 1 as shown in Figure 6(b).

In this paper, the scanning method is applied differently to one- and two-hop anchor nodes. If the normal node has two-hop anchor nodes, the values of cells, that are located within the data transmission radius r of each two-hop anchor node or outside a circle of radius $2r$ of each two-hop anchor node, are changed to 0 that means the normal node cannot be located in the cell. If the normal node has one-hop anchor nodes, the values of cells that are located outside of data transmission radius r are changed to 0. After scanning all of cells, the average coordinate of cells whose values are 1 is calculated as the estimated position of the normal node. Figure 6 shows an example of grid scanning. The ER formed in Figure 6 should be divided by many cells; however, we divide ER by 12 cells to understand it easily. The result of the grid scanning shows that the estimated position of a normal node N is N', and N' is used as the initial position of the normal node when correcting the position in Section 3.2.

Since sensor nodes are deployed in the sensor field randomly, there can be normal nodes without any information of anchor nodes. Such normal nodes wait until the neighbor normal nodes estimate and broadcast their position. Then, they estimate their position by the information from neighbor normal nodes. For the estimation by neighbor normal nodes, all normal nodes that obtain their position, broadcast their position information and their one-hop anchor nodes. The normal nodes, that do not have any position information of an anchor node, receive the information of neighbor normal nodes and mark one-hop anchor nodes of the neighbor

normal nodes as two-hop anchor nodes. They then perform the same ER construction and distributed grid scan as just explained steps.

3.2. Reducing Error Propagation (REP). The REP handles only the error propagation by neighbor positions with low precision (e.g., one case that two anchor nodes are located together, thus error propagation would be enlarged as shown Figure 4, or the other case by insufficient anchor nodes within two-hop). Through *hop progress calculation* and *repositioning*, REP corrects the position information of normal nodes by rotating the estimated positions centering on anchor nodes. Lastly in this subsection, we explain about decision method for whether to apply REP.

3.2.1. Hop Progress Calculation. To reduce error propagation, other information is necessary for adjusting position information. REP uses the information of anchor nodes that is multihop away from normal nodes. Normal nodes can obtain the information of anchor nodes multihop away when other normal nodes finishing to estimate the initial position and broadcasting their position to their neighbor normal nodes. When sending information of themselves to their neighbor sensor nodes after completing to estimate their initial position, normal nodes that have neighbor anchor nodes send not only their estimated position information but also the position information of an one-hop anchor node and the distance between the normal node and the one-hop anchor node. The anchor node information obtained first is used as the center of the rotation. Moreover, the center of the rotation of each normal node that does not have any neighbor anchor nodes is fixed for the center of the rotation of the neighbor normal node. The distance between the normal node and the center of the rotation is determined by hop progress (i.e., an estimated hop distance using probability based on the density of sensor nodes) [24]. The density of normal nodes means the average number of sensor nodes in the transmission radius of a sensor node. Hop progress can be derived from (1) [24, 25]. In (1), $E(R)$ is the estimated hop progress, E_C is the sensor node density, and r_0 is the transmission radius. Hop progress is calculated by only the sensor node density and the data transmission radius:

$$E(R) = \frac{2(E_C + 1)\sin\theta}{\pi {r_0}^2} \int_0^{r_0} l^2 e^{-((E_C+1)/\pi)\theta(1-(l^2/{r_0}^2))} dl. \quad (1)$$

Each normal node estimates the distance between itself and the center of the rotation using hop progress. Then, the normal node broadcasts the position information of the center of the rotation and hop progress to the position for neighbor normal nodes that do not have any information of anchor nodes. This broadcast can be efficient by adding the information of the center of the rotation to the broadcast packet after the grid-scan phase. Through this procedure, all normal nodes have position information of 1∼3 anchor nodes and hop progress to the anchor nodes as well as their initial position.

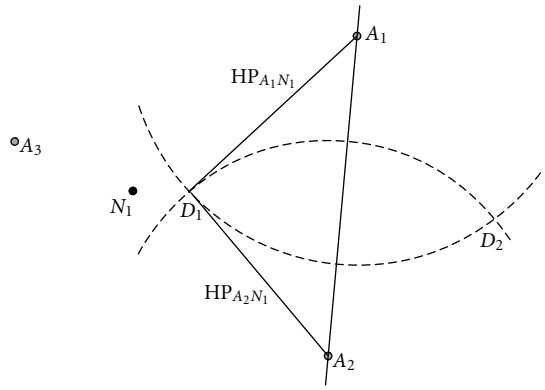
3.2.2. Repositioning. Each normal node that has the position information of three anchor nodes and hop progress to the anchor nodes corrects its initial position using its information obtained. Each normal node calculates two cross-points of two circles whose center point is the center of the rotation and radius is the hop progresses of two anchor nodes. These two cross-points are regarded as the position in which the normal node can be located. However, since the hop progresses are based on the probability, the accuracy of these two cross-points are low. Thus, these two points are used not to correct the initial position directly, but to decide the direction of the rotation. The direction of the rotation is decided as a cross-point of two circles that is nearer by the third anchor node that is not used to make two circles. The reason of the selection between two cross-points is that the cross-point far from the third anchor node cannot receive the information of the third anchor node.

To correct the initial position of a normal node by rotation, the angle of the rotation is necessary. As knowing the initial position and the center and the direction of the rotation, the normal node can obtain the angle of the rotation. Figure 7 shows an example of the center and the angle of the rotation, and the angle of the rotation can be obtained by (2):

$$\cos\theta = \frac{d(A_1, D)^2 + d(A_1, {N_1}')^2 - d({N_1}', D)^2}{2 \cdot d(A_1, D) \cdot d(A_1, {N_1}')}. \quad (2)$$

In (2), $d(a, b)$ is the Euclidean distance between the point a and b. The value of $\sin\theta$ can be a positive or negative value, so the normal node should decide the presence of the negative sign. When the coordinate of A_1 is (x_{A_1}, y_{A_1}), the coordinate of ${N_1}'$ is $(x_{{N_1}'}, y_{{N_1}'})$, and the coordinate of the rotation direction D is (x_D, y_D), $\sin\theta$ is calculated in (3) by the relation between the line $A_1{N_1}'$ and the position D:

$$\sin\theta = \begin{cases} \sqrt{1-(\cos\theta)^2}, & \text{if } \left(x_{{N_1}'} - x_{A_1} \geq 0,\ (x_{{N_1}'} - x_{A_1})(y_D - y_{A_1}) \geq (y_{{N_1}'} - y_{A_1})(x_D - x_{A_1})\right) \\ & \text{or } (x_{{N_1}'} - x_{A_1} < 0,\ (x_{{N_1}'} - x_{A_1})(y_D - y_{A_1}) < (y_{{N_1}'} - y_{A_1})(x_D - x_{A_1})) \\ -\sqrt{1-(\cos\theta)^2}, & \text{otherwise.} \end{cases} \quad (3)$$



A_i: Anchor node
N_i: Normal node
D_i: Intersection point with two circles
$HP_{A_iN_j}$: Hop progress between A_i and N_j

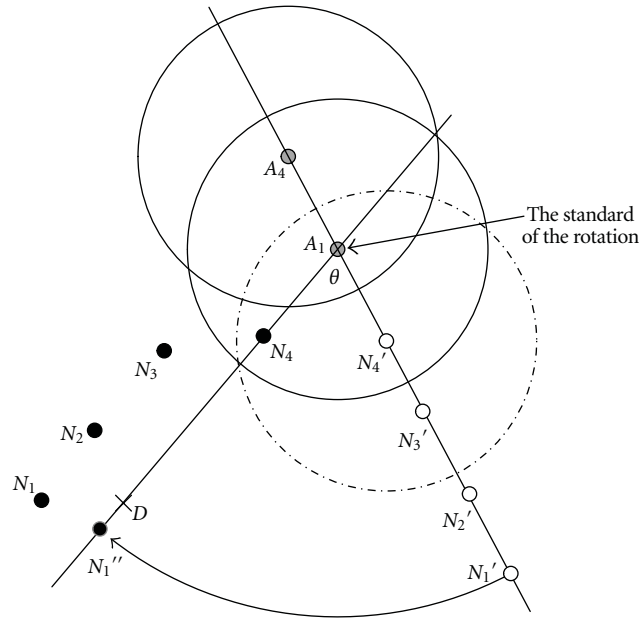
Figure 7: Deciding the direction of the rotation with the information of three anchor nodes.

Then, the coordinate of $N_1{}''$, $(x_{N_1''}, y_{N_1''})$ that is the rotatory translation of $N_1{}'$ with the angle of rotation θ is obtained in (4):

$$\begin{pmatrix} x_{N_1''} \\ y_{N_1''} \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_{N_1'} - x_{A_1} \\ y_{N_1'} - y_{A_1} \end{pmatrix} + \begin{pmatrix} x_{A_1} \\ y_{A_1} \end{pmatrix}. \quad (4)$$

In Figure 8, $N_1{}''$ that is the corrected position by rotatory translation is closer to N_1, that is the real position of the normal node, than the initial position $N_1{}'$. The normal nodes that complete REP broadcast their estimated position information to apply REP to the normal nodes that have one or two anchor node information.

3.2.3. Decision for Applying REP. REP does not always reduce the estimation error. Even though REP is applied to amend the initial estimated positions of normal nodes, the error of the localization can be increased when the initial estimated positions are relatively correct. Therefore, whether to apply REP is decided for more accurate localization. Each normal node compares the variations of estimated positions before and after applying REP. The variation of the estimated position means the degree of error. The bigger the region that a normal node can be located in is, the bigger the variation is. When the variation of estimate position after applying REP is smaller than that before applying REP, REP is applied to the initial estimated position. On the contrary, the initial estimated position is decided as the final estimated position when the variation of estimate position after applying REP is bigger than that before applying REP. Generally, normal nodes that are close to anchor nodes have low variance for the estimated position, since the degree of error propagation is relatively small when normal nodes that are close to anchor nodes estimate their initial position.



A_i: Anchor node
N_i: Normal node
N_i': Initial estimated position
N_i'': Final estimated position
θ: The angle of the rotation
D: The direction of the rotation

Figure 8: Amending the position with the initial estimated position and the direction.

4. Performance Evaluation

This section handles the performance results of ADL and ADL with REP. To show the degree of performance improvement, ADL and ADL with REP are compared to DRLS, a previous range-free localization scheme. The earlier part of this section presents the simulation environment, and the metrics and parameters of performance evaluation. The later part of this section presents the localization accuracy and message transmission cost of DRLS, ADL, and ADL with REP.

4.1. Methodology. The simulator in this section is realized using JAVA. Data transmission radiuses of all sensor nodes are same as r in the simulation of ADL and ADL with REP. The unit disk model [26] is used to transmit data. That is, a sensor node transmits data successfully to its neighbor sensor nodes that are located within r. The same sensor nodes are deployed randomly on the sensor field whose size is $10r \times 10r$. A few anchor nodes that equip GPS know their real position. All sensor nodes are the same, except for the existence of GPS. The collision that occurs during message transmission for the position estimation is not considered. The cell size of the grid scan phase in ADL is fixed to $0.1r \times 0.1r$. All simulation results are based on 100 repetitions.

The metrics for the performance evaluation are localization accuracy, mean squared error (MSE), and message transmission cost.

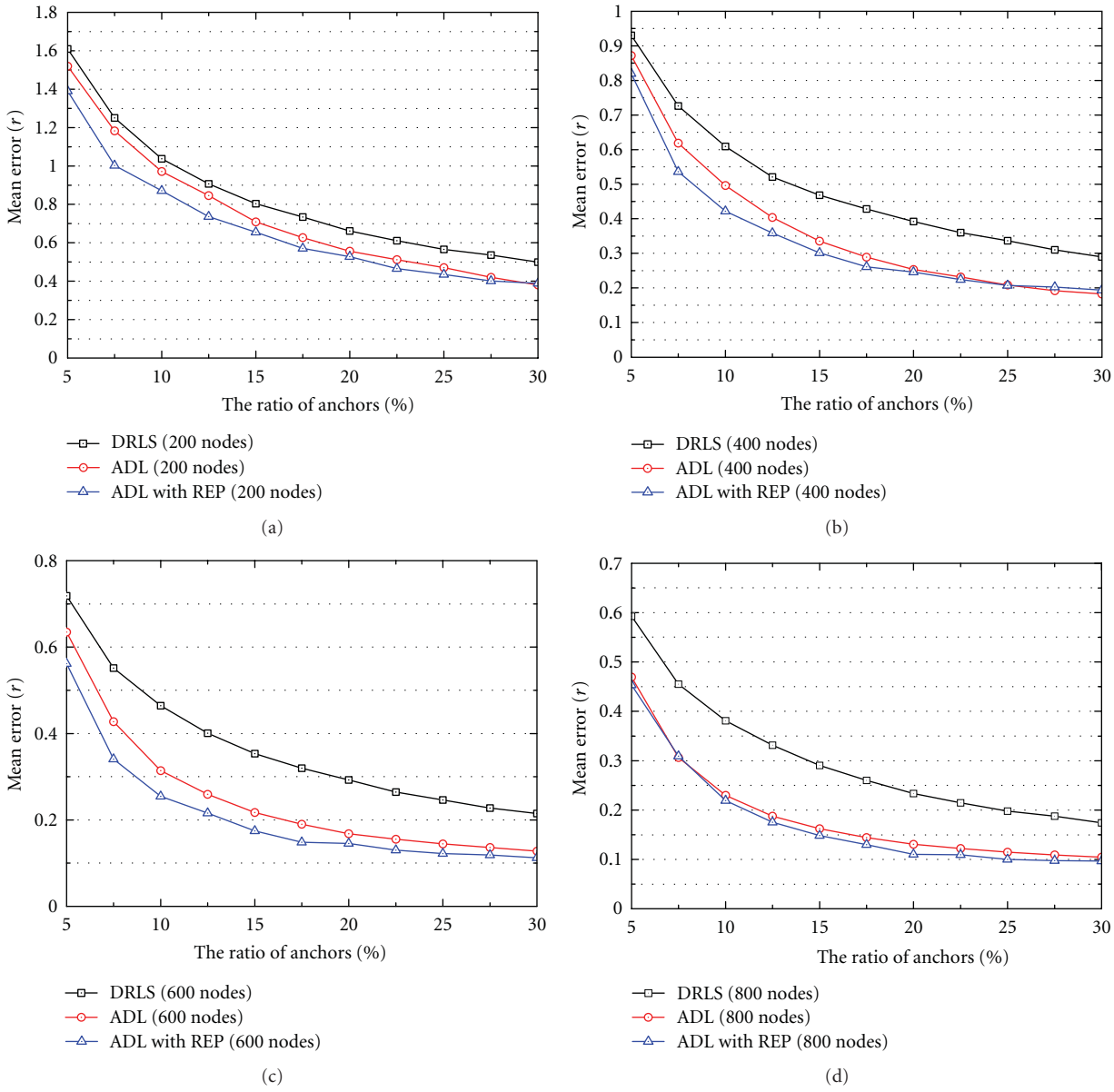


Figure 9: Mean error versus anchor ratio: (a) the number of sensors is 200; (b) 400; (c) 600; (d) 800.

(1) *Localization accuracy.* Localization accuracy is defined as the average of the Euclidean distance between the real positions and the estimated positions of all normal nodes. The data transmission radius r is the basic unit of localization accuracy. Localization accuracy denotes how accurately normal nodes estimate their positions.

(2) *Mean squared error (MSE).* MSE for the estimated position of normal nodes denotes the degree of variance of the estimated position of normal nodes.

(3) *Message transmission cost.* Message transmission cost is the number of messages used for normal nodes to obtain the information of anchor nodes.

The parameters for the simulation are the density of sensor nodes and the ratio of anchor nodes. The density of the sensor nodes is diversified by the change of the number of sensor nodes, and the number of anchor nodes is changed for the density of sensor nodes to be varied over the entire sensor field. Since the ratio of the anchor nodes is controlled by the fixed number of sensor nodes deployed in the sensor field, the number of normal nodes is decreased if the number of anchor nodes is increased.

4.2. Average Error and Mean Squared Error with Different Anchor Ratio. This section analyzes the localization accuracy and MSE of ADL and ADL with REP versus the ratio of anchor nodes. The data transmission radius of each sensor node is r. The simulation is performed against various numbers of

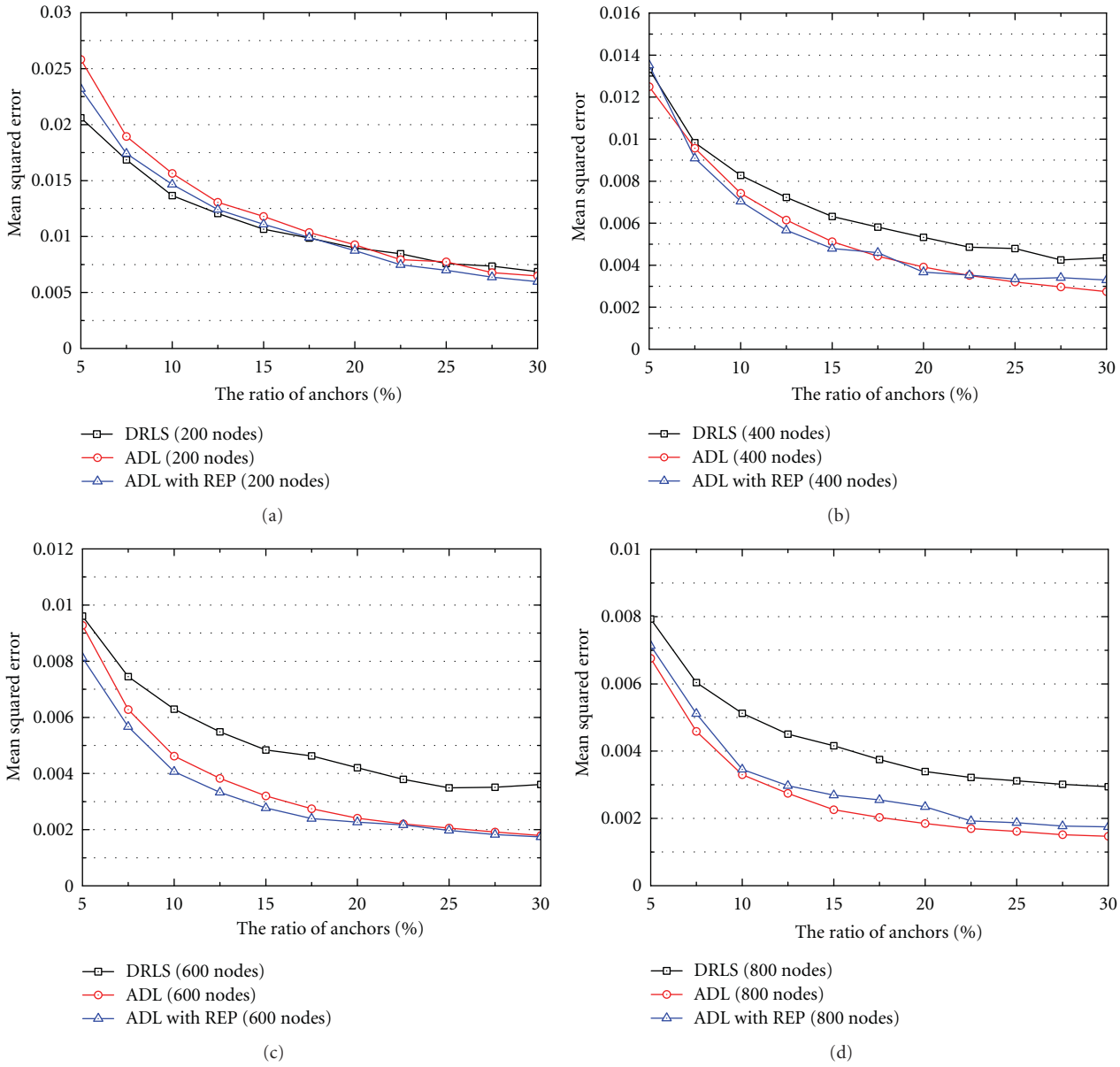


Figure 10: MSE versus anchor ratio: (a) the number of sensors is 200; (b) 400; (c) 600; (d) 800.

sensor nodes, such as 200, 400, 600, and 800 in the sensor field whose size is $10r \times 10r$, in addition, the ratio of anchor nodes is changed from 5 to 30% for each number of sensor nodes.

Figure 9 shows the mean error of the estimated position of DRLS, ADL, and ADL with REP in accordance with the variety of the ratio of anchor nodes. Localization accuracy of ADL with REP is the highest, and that of DRLS is the least, since ADL increases the localization accuracy by estimating the position of normal nodes without the vector-based refinement of DRLS. ADL with REP amends the initial estimated position of normal nodes in ADL by applying REP. The lower the ratio of anchor nodes is, the fewer the amount of the position information of anchor nodes is; therefore, localization accuracy is decreased. When the number of sensor nodes is low, localization accuracy is relatively low due to the hole problem. Moreover, when the number of the sensor nodes is high, localization accuracy is increased due to the increase in the information that a normal node can obtain. When there are 800 sensor nodes, the difference of localization accuracy between ADL and ADL with REP is little due to the fact that the initial estimated positions are already sufficiently accurate, and the estimation error can be increased by REP. The simulation results show that ADL with REP has a maximum of 18% higher accuracy compared to ADL and a maximum of 30% higher accuracy compared to DRLS, when there are 400 sensor nodes.

Figure 10 shows MSE of DRLS, ADL, and ADL with REP versus the ratio of anchor nodes. Except for the cases with 200 and 800 sensor nodes, ADL with REP has the lowest MSE,

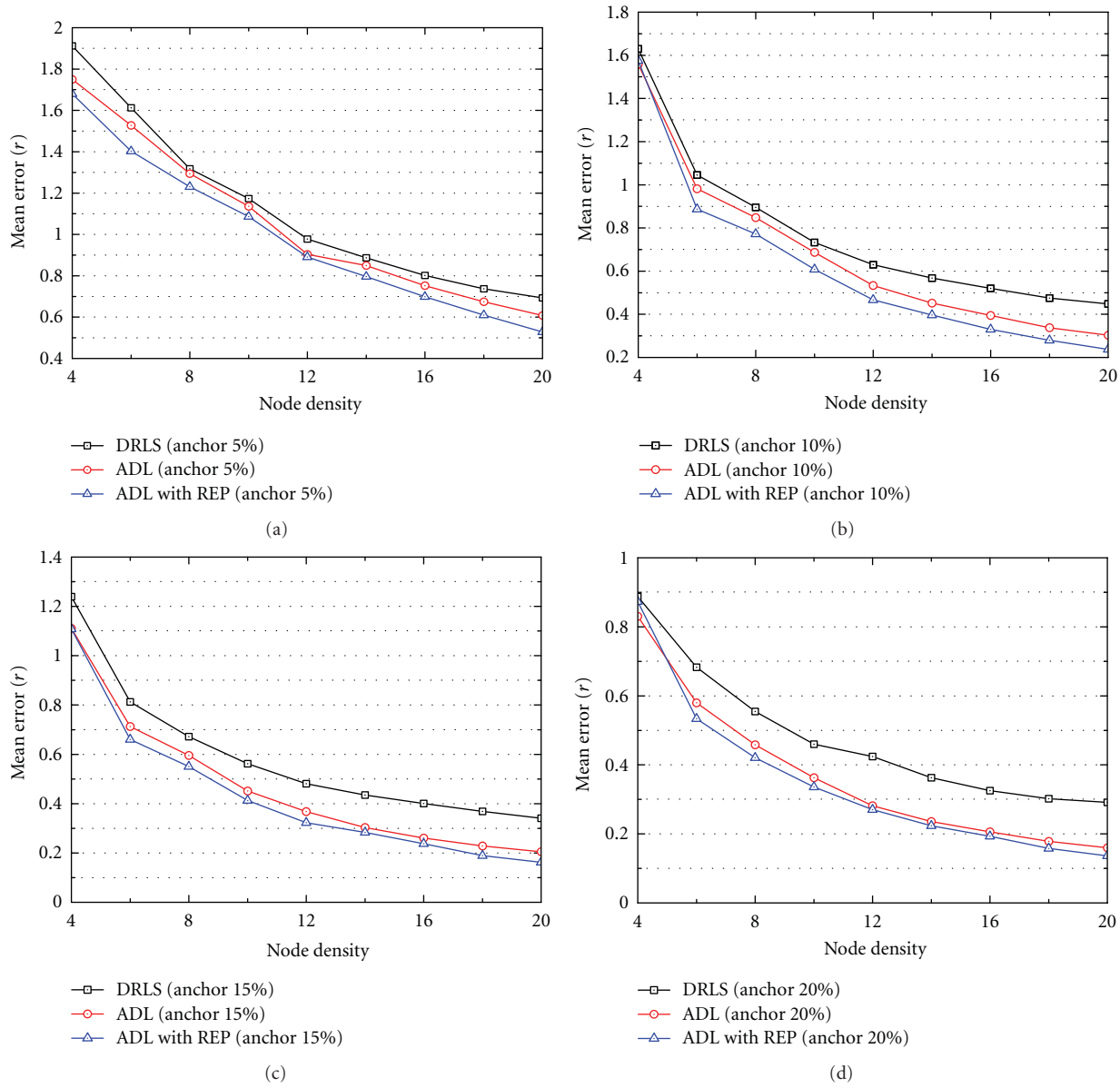


Figure 11: Mean error versus node density: (a) anchor ratio is 5%; (b) 10%; (c) 15%; (d) 20%.

and DRLS has the highest MSE since ADL with REP has the least variance of the estimated position by reducing error propagation. For 200 sensor nodes, MSE of DRLS is the lowest when the ratio of the anchor nodes is low, since the number of normal and anchor nodes is small and thus each normal node cannot receive the sufficient position information. When there are 800 sensor nodes, the MSE of ADL of REP is high, since the normal nodes, whose positions are already accurate, amend their positions by REP, so that the position estimation error is increased. The simulation results show that ADL with REP has a maximum of 7% lower MSE compared to ADL and a maximum of 22% lower MSE compared to DRLS when there are 400 sensor nodes.

4.3. Average Error and Mean Squared Error with Different Sensor Node Density. This section analyzes the localization accuracy and MSE with various densities of sensor nodes. The density of sensor nodes is defined as the average number of sensor nodes within the data transmission radius of a sensor node. When the data transmission radius is r, if 200 sensor nodes are deployed in the sensor field whose size is $10r \times 10r$, the number of sensor nodes in the data transmission radius is 6.28. The simulation is performed with anchor ratios of 5, 10, 15, and 20%, and the density of sensor nodes is changed from 4 to 20 in each density of anchor nodes.

Figure 11 shows the localization accuracy versus the density of sensor nodes. Mean error of ADL with REP is the smallest and that of DRLS is the largest, since ADL increases localization accuracy by estimating the position of normal nodes without the vector-based refinement of DRLS. ADL with REP amends the initial estimated position of normal

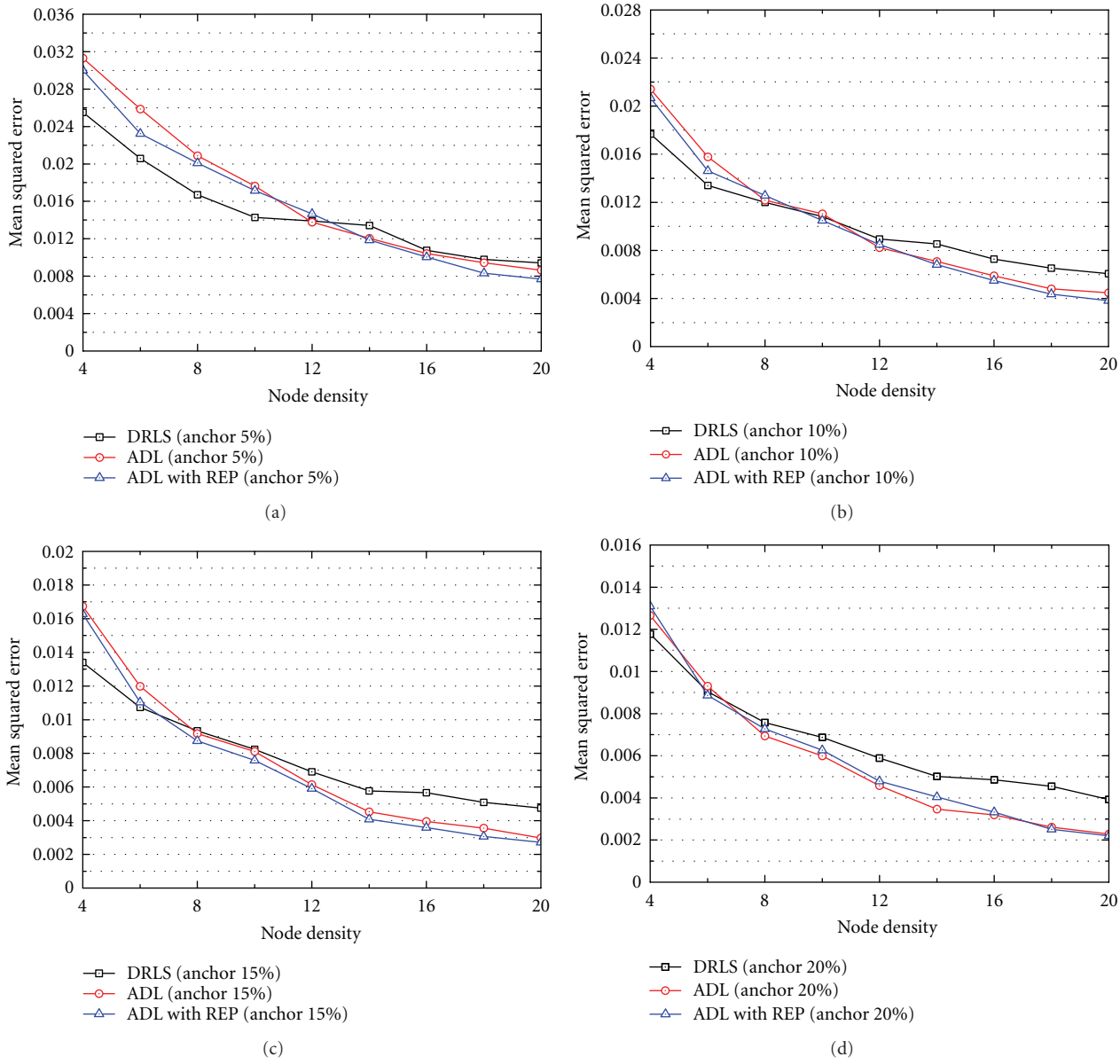


Figure 12: MSE versus node density: (a) anchor ratio is 5%; (b) 10%; (c) 15%; (d) 20%.

nodes in ADL by applying REP. When the ratio of anchor nodes is 20% and the density of sensor nodes is 4, mean error of ADL is lower than that of ADL with REP since the number of both normal nodes and anchor nodes is small and thus accurate position estimation is impossible. The simulation results show that ADL with REP has a maximum of 13% higher accuracy compared to ADL and a maximum of 25% higher accuracy compared to DRLS when the density of anchor nodes is 10%.

Figure 12 shows the MSE versus the various densities of sensor nodes. Due to the fact that more accurate position estimation is possible with many sensor nodes, the higher the density of sensor nodes is, the lower the MSE is. When the density of sensor nodes is low, MSE of DRLS is the lowest since normal nodes cannot receive sufficient position information from the small number of normal and anchor nodes. When the ratio of anchor nodes is 20%, MSE of ADL is lower than that of ADL with REP, since normal nodes whose initial estimated positions are sufficiently accurate from many anchor nodes amend their positions by REP so that MSE is increased. When the ratio of anchor nodes is 10%, the simulation results show that ADL with REP has a maximum of 8% lower MSE compared to ADL, and a maximum of 28% lower MSE compared to DRLS with an environment whose node density exceeds 12.

4.4. Message Transmission Cost. This section analyzes the message transmission cost for position estimation by ADL and ADL with REP. This cost is defined as the number of messages exchanged. The simulation environment is changed with the ratio of anchor nodes from 5 to 30% and the density of sensor nodes from 4 to 20.

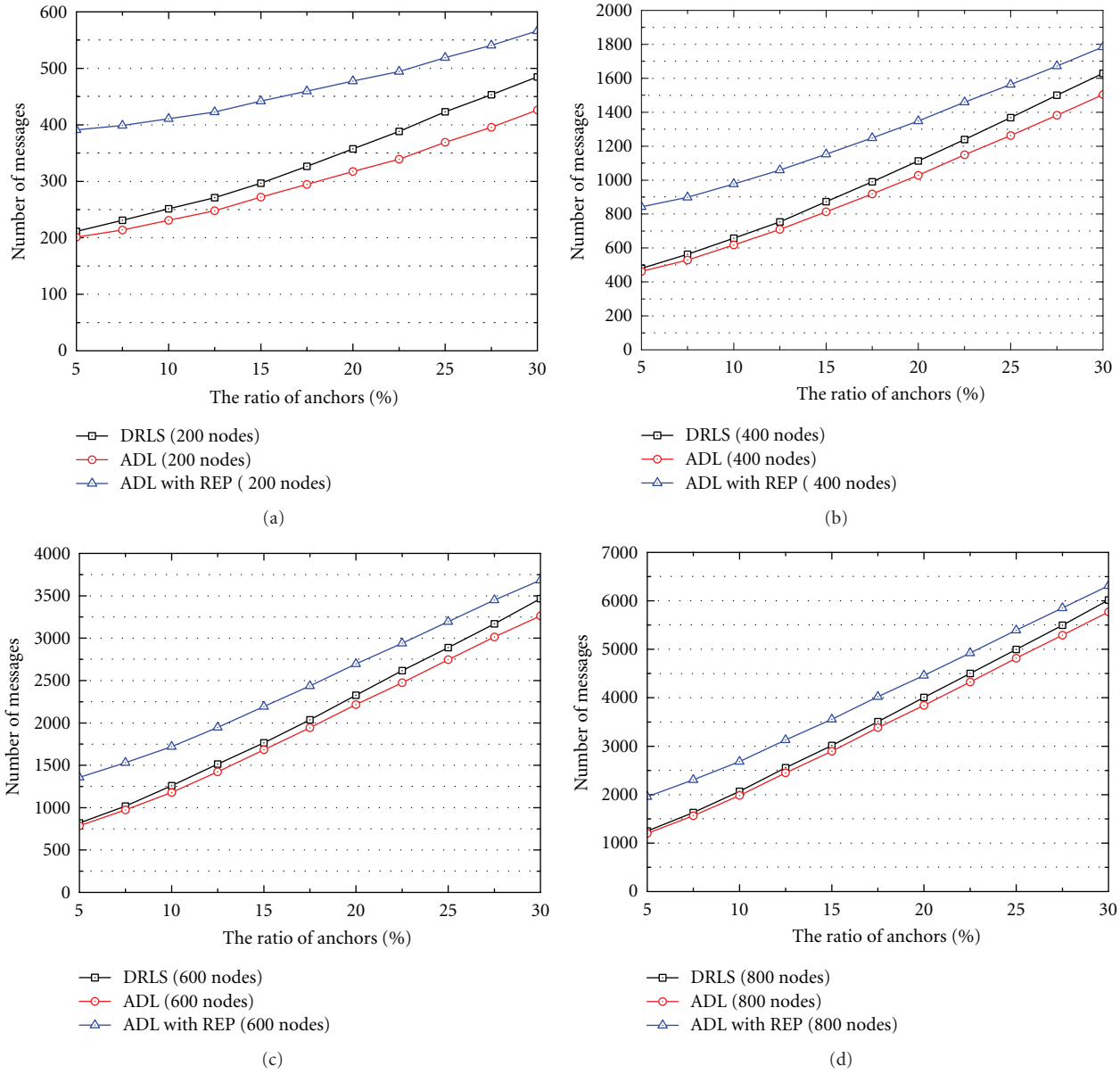


Figure 13: Number of messages versus anchor ratio: (a) the number of sensors is 200; (b) 400; (c) 600; (d) 800.

Figure 13 shows the message transmission cost against various ratios of anchor nodes. Due to the fact that two-hop flooding is performed based on the anchor nodes in DRLS and based on normal nodes in ADL, the message transmission cost of ADL is the lowest. That is, each anchor node in DRLS exchanges its position information with anchor nodes that are located within two-hop distance, and provides all of information to its neighbor normal nodes. Even though normal nodes in DRLS obtain some position information during the information exchange between anchor nodes, anchor nodes after finishing the two-hop flooding transmit the information to normal nodes. Due to such unnecessary transmissions of DRLS, ADL outperforms DRLS a little from the viewpoint of transmission cost. The higher data transmission cost of ADL with REP is due to each normal node broadcasting the information of their center of the rotation additionally after applying REP. Generally, DRLS transmits more messages as many as the number of anchor nodes than ADL, and ADL with REP transmits more messages as many as the number of normal nodes than ADL. The simulation results show that the message transmission cost of ADL with REP is a maximum of 45% higher than ADL when the ratio of anchor nodes is low, and a maximum of 35% higher than DRLS when the number of sensor nodes is 400 and the ratio of anchor node is 10%.

Figure 14 shows the message transmission cost against various densities of sensor nodes. Like Figure 13, the message transmission cost of ADL is the lowest and that of ADL with REP is the highest since two-hop flooding of ADL is more efficient than that of DRLS. ADL with REP makes normal

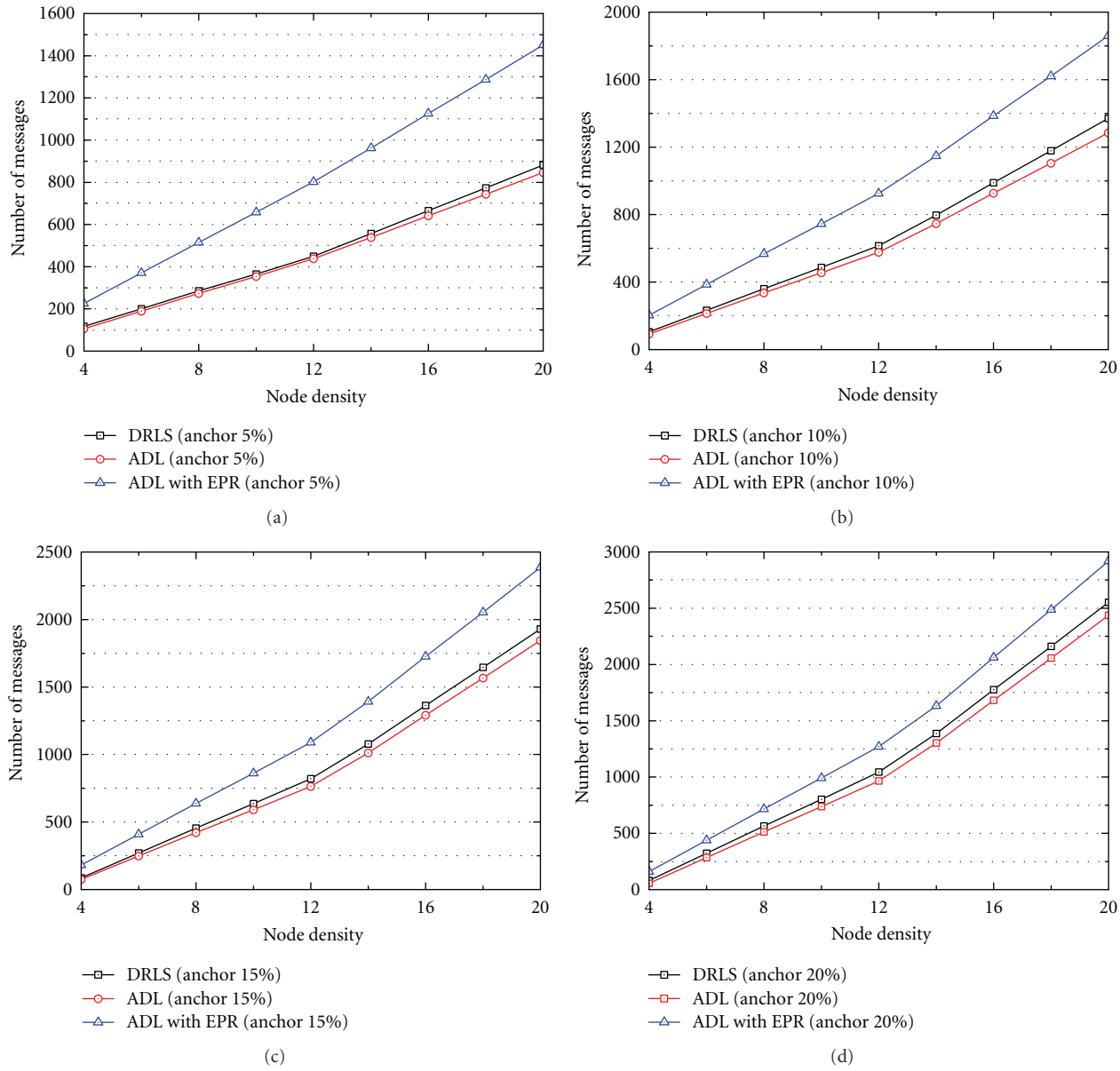


Figure 14: Number of messages versus node density: (a) anchor ratio is 5%; (b) 10%; (c) 15%; (d) 20%.

nodes to amend their initial estimated positions by additional message transmissions. When the ratio of anchor nodes is low such as 5%, in ADL with REP, normal nodes transmit their center of the rotation to their neighbor normal nodes additionally after applying REP, so there is a big gap between transmission costs of ADL and ADL with REP. The simulation results show that the message transmission cost of ADL with REP is a maximum of 45% higher than ADL and that of ADL is a maximum of 7% lower than DRLS when the ratio of anchor nodes is 10% and the density of sensor nodes is 10.

5. Conclusions

The paper has proposed ADL that is a distributed localization scheme to increase the position accuracy and reduce the message transmission cost, and ADL with REP that is a scheme to reduce the error propagation. In ADL, each normal node assumes an ER using anchor nodes within its transmission radius that is obtained by two-hop flooding, and initial position is estimated by grid scanning in the ER. However, since the initial estimated positions are affected by the error propagation in the environment with the small number of anchor nodes, initial estimated positions should be amended. The ADL with REP makes normal nodes to estimate their position more accurately by rotating the initial estimated positions whose centers are anchor nodes.

The simulation results with the various ratios of anchor nodes and the various densities of sensor nodes showed that localization accuracy of ADL is 12% higher and the data transmission cost is 10% lower than DRLS. In addition,

localization accuracy of ADL with REP is 30% higher, but the data transmission cost is 35% higher than DRLS. Due to the fact that the data transmission cost of ADL with REP is high, we should decide a scheme among ADL and ADL with REP according to the application. We will enhance the proposed localization in order that the case whereby the amended position is more inaccurate than the initial estimated position by applying REP is avoided by the statistical analysis.

Acknowledgment

This research was supported in part by MKE and MEST, Korean government, under ITRC NIPA-2012-(H0301-12-3001), NGICDP(2011-0020517) and PRCP(2011-0018397) through NRF, respectively.

References

[1] 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.

[2] G. Djuknic and R. Richton, "Geolocation and assisted GPS," *IEEE Computer*, vol. 34, pp. 123–125, 2001.

[3] C. Wang and X. Li, "Sensor localization under limited measurement capabilities," *IEEE Network*, vol. 21, no. 3, pp. 16–23, 2007.

[4] A. Boukerche, H. A. B. F. Oliveira, E. F. Nakamura, and A. A. F. Loureiro, "Localization systems for wireless sensor networks," *IEEE Wireless Communications*, vol. 14, no. 6, pp. 6–12, 2007.

[5] G. J. Pottie and W. J. Kaiser, "Wireless integrated network sensors," *Communications of the ACM*, vol. 43, no. 5, pp. 51–58, 2000.

[6] N. B. Priyantha, A. Chakraborty, and H. Balakrishnan, "Cricket location-support system," in *Proceedings of the6th Annual International Conference on Mobile Computing and Networking (MOBICOM'00)*, pp. 32–43, August 2000.

[7] A. Savvides, C. 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.

[8] J. Hightower and G. Borriello, "Location systems for ubiquitous computing," *Computer*, vol. 34, no. 8, pp. 57–66, 2001.

[9] T. Cong and I. Koo, "An RSS-based localization scheme using direction calibration and reliability factor information for wireless sensor networks," *KSII Transactions on Internet and Information Systems*, vol. 4, no. 1, pp. 45–61, 2010.

[10] N. Bulusu, J. Heidemann, and D. Estrin, "GPS-less low-cost outdoor localization for very small devices," *IEEE Personal Communications*, vol. 7, no. 5, pp. 28–34, 2000.

[11] L. Doherty, K. S. J. Pister, and L. El Ghaoui, "Convex position estimation in wireless sensor networks," in *Proceedings of the 20th Annual Joint Conference of the IEEE Computer and Communications Societies*, pp. 1655–1663, April 2001.

[12] D. Niculescu and B. Nath, "Ad Hoc Positioning System (APS)," in *Proceedings of the IEEE Global Communications Conference (GLOBECOM '01)*, pp. 2926–2931, November 2001.

[13] J. P. Sheu, J. M. Li, and C. S. Hsu, "A distributed location estimating algorithm for wireless sensor networks," in *Proceedings of the IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing*, pp. 218–225, June 2006.

[14] L. Hu and D. Evans, "Localization for mobile sensor networks," in *Proceedings of the 10th Annual International Conference on Mobile Computing and Networking (MobiCom'04)*, pp. 45–57, October 2004.

[15] M. Rudafshani and S. Datta, "Localization in wireless sensor networks," in *Proceedings of the 6th International Symposium on Information Processing in Sensor Networks (IPSN '07)*, pp. 51–60, April 2007.

[16] R. Stoleru, P. Vicaire, T. He, and J. A. Stankovic, "StarDust: a flexible architecture for passive localization in wireless sensor networks," in *Proceedings of the 4th International Conference on Embedded Networked Sensor Systems (SenSys'06)*, pp. 57–70, 2006.

[17] A. Baggio and K. Langendoen, "Monte-Carlo localization for mobile wireless sensor networks," in *Proceedings of the 2nd International Conference on Mobile Ad-hoc and Sensor Networks (MSN'06)*, pp. 317–328, December 2006.

[18] T. He, C. Huang, B. M. Blum, J. A. Stankovic, and T. Abdelzaher, "Range-free localization schemes for large scale sensor networks," in *Proceedings of the 9th Annual International Conference on Mobile Computing and Networking, (MobiCom'03)*, pp. 81–95, September 2003.

[19] P. Liu, X. Zhang, S. Tian, Z. Zhao, and P. Sun, "A novel virtual anchor node-based localization algorithm for wireless sensor networks," in *Proceedings of the 6th International Conference on Networking, (ICN'07)*, April 2007.

[20] H. Mirehrahim and M. Dehghan, "Monte Carlo localization of mobile sensor networks using the position information of neighbor nodes," in *Proceedings of the Springer ADHOC-NOW*, pp. 270–283, August 2009.

[21] J. P. Sheu, P. C. Chen, and C. S. Hsu, "A distributed localization scheme for wireless sensor networks with improved grid-scan and vector-based refinement," *IEEE Transactions on Mobile Computing*, vol. 7, no. 9, Article ID 4468707, pp. 1110–1123, 2008.

[22] T. Kim, M. Shon, W. Choi, M. Song, and H. Choo, "Low-cost two-hop anchor node-based distributed range-free localization in wireless sensor networks," in *Proceedings of the ICCSA*, pp. 165–174, March 2010.

[23] T. Kim, M. Shon, M. Kim, D. Kim, and H. Choo, "Reducing error propagation on anchor node-based distributed localization in wireless sensor networks," in *Proceedings of the International Conference Ubiquitous Computing and Multimedia Applications (UCMA '11)*, pp. 22–28, April 2011.

[24] Y. Wang, X. Wang, D. Wang, and D. P. Agrawal, "Range-free localization using expected hop progress in wireless sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 20, no. 10, pp. 1540–1552, 2009.

[25] X. Wang, *QoS issues and QoS constrained design of wireless sensor network*, Ph.D. dissertation, University of Cincinnati, 2006.

[26] B. N. Clark, C. J. Colbourn, and D. S. Johnson, "Unit disk graphs," *Discrete Mathematics*, vol. 86, no. 1-3, pp. 165–177, 1990.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 467124, 12 pages
doi:10.1155/2012/467124

Research Article

Networked Electronic Equipments Using the IEEE 1451 Standard—VisioWay: A Case Study in the ITS Area

F. Barrero, S. L. Toral, M. Vargas, and J. Becerra

Escuela Técnica Superior de Ingeniería, Universidad de Sevilla, Avenida Camino de los Descubrimientos s/n, 41092 Sevilla, Spain

Correspondence should be addressed to F. Barrero, fbarrero@esi.us.es

Received 16 September 2011; Accepted 6 February 2012

Academic Editor: Tai Hoon Kim

The concept of Intelligent Transportation Systems (ITSs) has been recently introduced to define modern embedded systems with enhanced digital connectivity, combining people, vehicles, and public infrastructure. The smart transducer concept, on the other hand, has been established by the IEEE 1451 standard to simplify the scalability of networked electronic equipments. The synergy of both concepts will establish a new paradigm in the near future of the ITS area. The purpose of this paper is to analyze the integration of electronic equipments into intelligent road-traffic management systems by using the smart transducer concept. An automated video processing sensor for road-traffic monitoring applications is integrated into an ITS network as a case study. The impact of the IEEE 1451 standard in the development and performance of ITS equipments is analyzed through its application to this video-based system, commercialized under the name VisioWay.

1. Introduction

Recent progress in semiconductor technology and microprocessor architectures, on the one hand, and the ubiquity of telecommunication networks, on the other, favour the growth of a novel market context for sensors and actuators [1, 2]. Sensors integrated into structures—coupled with the efficient delivery of sensed information—provide new services and benefits to society [3–5]. The actual implementation of such systems occurs on local area networks (LANs) that look like hierarchical networks with distributed computing capabilities. Using this technology, sensors and actuators provide a significant improvement in performance and effectiveness through the distribution of their resources.

It is interesting to link the development of such distributed data acquisition systems to the state of the art and future perspectives in the ITS field, which has attracted much attention in recent years [6]. ITS equipments are nowadays evolving into complex systems, allowing vehicle-to-vehicle (V2V) communication and cooperation [7, 8], or vehicle-to-infrastructure (V2I) interaction [9]. The public infrastructure (traffic lights, streetlamps or traffic signs) is adapting to these technological advances. Some examples are prompting or advisory message systems, informing the drivers on particular traffic conditions—including alerts on imminent dangerous situations—or automated traffic-data acquisition systems transmitting relevant data to traffic control centers—traffic flow, per-lane vehicle counting, queue-length estimation, congestion level, estimation of journey times, red light runner detection, and so forth [7, 10, 11].

However, there are still some handicaps to this widespread use of electronic equipments in the ITS field: risk of breakage or connector failures in today wired systems, limited scalability in current networked systems bound to vendor-specific protocols, and so forth.

Wireless technologies are the best means to overcome these limitations, contributing for instance to a substantial ease of installation, to an increase of scalability [12] or to alleviate maintenance of physical links. Nowadays, an ideal wireless network should have some "smart" features, such as being highly scalable, software programmable, energy efficient, inexpensive, and hopefully plug-and-play compliant; on top of that, it has to be reliable and accurate, requiring a very low maintenance effort over the long term.

Recently, one of the most relevant milestones in this context is the introduction of the concept "smart transducer," ac cording to the IEEE 1451 standard [13, 14]. IEEE 1451 provides an open platform for the development of scalable

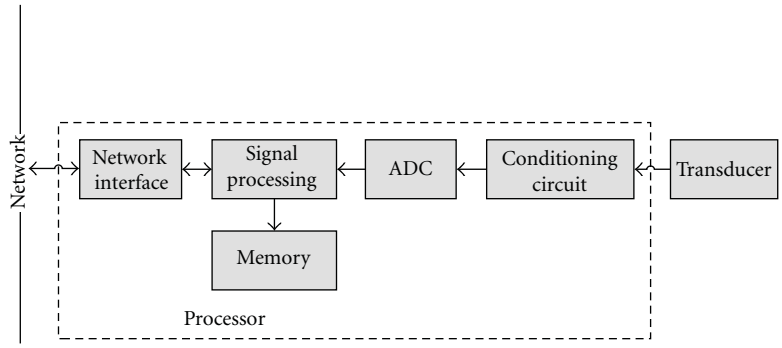


Figure 1: Smart transducer architecture.

networked electronic modules, using different physical links. The standard offers the ability to simplify the connectivity of electronic equipments, allowing the "plug and play" capability by means of the physical-link abstraction. These, among others, are some expected, interesting features in any IEEE1451-compliant equipment, leading to increased scalability and less expensive maintenance.

This paper analyzes the utility of IEEE 1451 for the development of ITS equipments. An IEEE 1451-compliant prototype for urban road-traffic monitoring is described and analyzed as a case study. The new prototype is built upon an existing intelligent video sensor, called VisioWay, which was developed for road-traffic monitoring applications (http://www.visioway.com/). VisioWay has been envisaged as part of the road-traffic infrastructure, for traffic monitoring and data collection. The relevant information is transmitted to the traffic control centers, to be used for traffic-control assessment or road-traffic management.

The paper is organized as follows. First, a brief overview of the IEEE 1451 standard and the smart transducer concept is given in Section 2. Next, Section 3 describes the preexisting VisioWay system and its use in road-traffic monitoring applications. Then, the integration of the IEEE 1451 standard into the VisioWay equipment is analyzed in Section 4, along with the description of the experimental setup. Finally, the conclusions are depicted in Section 5.

2. The Smart Transducer Concept: IEEE 1451 Standard Overview

The purpose of the IEEE 1451 standard is to define a set of common interfaces, functionalities, and commands for connecting sensors and actuators to microprocessor-based systems, instruments, and field networks in a network-independent fashion [13]. The main goal of this family of standards is developing network- and vendor-independent transducer interfaces, allowing electronic transducers to be replaced and/or moved with the minimum effort. Other desirable features include suppressing manual configuration steps. The Transducer Electronic Data Sheet (TEDS) concept is developed as an electronic datasheet that remains with the transducer during normal operation, allowing the real-time access to data and calibration details. Consequently, the IEEE 1451 family of standards provides an open platform for the development of networked electronic equipments.

Transducers that are 1451-compliant are also called smart transducers. Considering the purpose of the IEEE 1451 standard, a smart transducer should include certain elements like a physical transducer, a processor/memory core, and a local network interface through a microprocessor device. The transducer senses a given physical magnitude, generating the corresponding electrical signal. This signal is converted into digital values to be processed by the microcontroller, which delivers the resulting data to a local network. Only the selected or preconfigured data are accessible through the local network. Figure 1 depicts the architecture of a conventional smart transducer.

Two major components are defined, the Network Capable Application Processor (NCAP) and the Transducer Interface Module (TIM). The NCAP represents a network node that performs application processing and network communication functions, while the TIM consists of a transducer system, used for signal conditioning and data conversion, including up to 255 sensors and actuators. Figure 2 shows the overall structure for the standard IEEE 1451.

Eight separate standards are included in the IEEE 1451 family, from the IEEE 1451.0 to the IEEE 1451.5, the IEEE P1451.6, and the IEEE 1451.7.

The IEEE 1451.0 is one of the newest members of this family of standards. It defines a common set of commands, functionalities, and the format of the exchanged messages between the TIM and NCAP. The standard also provides the structure of the Transducer Electronic Data Sheets (TEDSs). The set of commands—independent of the physical communications media—includes the basic functions for reading and writing the transducers data, reading and writing TEDS, and sending configuration, control, and operation commands to the TIM. The aim of IEEE 1451.0 standard is to guarantee compatibility across the IEEE 1451 family.

The IEEE 1451.1 standard defines an information model for the communication between different NCAPs or between NCAPs and other systems [14]. It defines the NCAP as the main interface element between transducers and the network. The IEEE 1451 family of standards presents three possible protocols that can be indistinctly used to access electronic devices from the Internet: the Hyper Text Transfer Protocol (HTTP) outlined in the IEEE 1451.0 standard,

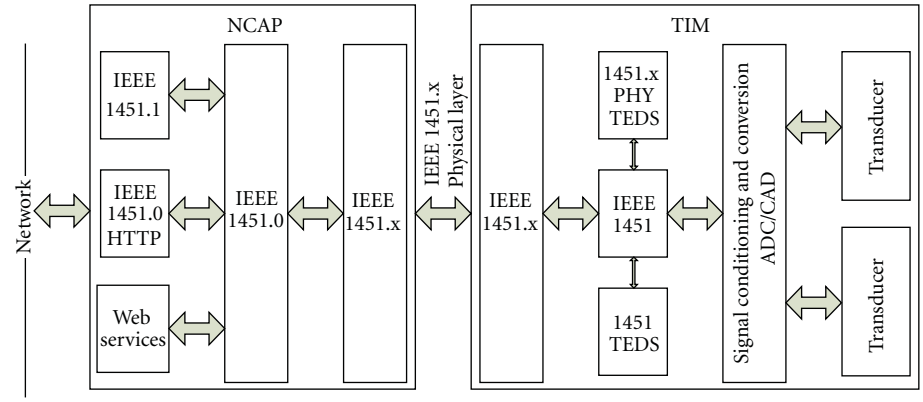


Figure 2: IEEE 1451 architecture overview.

Table 1: Generic format for TEDS.

Field	Description	Type	No. octets
0	TEDS length	UInt32	4
1 to N	Datablock	Variable	Variable
$N + 1$	Checksum	UInt16	2

an object and data model introduced in the IEEE 1451.1-1999 standard, and the proposed Smart Transducer Web Services or STWS derived from the Web Services Description Language or WSDL [15, 16]. The simplest web access is the HTTP protocol provided by the IEEE 1451.0 standard. The IEEE 1451.0 standard introduces an HTTP API that can be used to access the NCAP over TCP/IP networks. This HTTP protocol focuses on accessing transducer data and TEDS through the HTTP 1.1 protocol. Users send an HTTP request to a HTTP server running on the NCAP, receiving response from the server in Hyper Text Markup Language (HTML), eXtended Markup Language (XML), or Text format. HTTP and IEEE 1451.0 by themselves do not provide any means to discover the nodes in the network; nevertheless, there are several off-the-shelf network discovery tools that make use of different protocols that can be used for that matter.

Different physical interfaces are allowed between NCAP and TIMs, being one of the most interesting characteristics of the standard. For instance, the IEEE 1451.2 standard, first approved in 1997, specified a ten-wire point-to-point interface called Transducer Independent Interface or TII and the related protocol [17]. It defines a unique TIM named STIM or Smart Transducer Interface Module. However, a revision of this standard is expected due to the lack of manufacturer's interest and its incompatibility with the IEEE 1451.0 standard. This future revision will allow the support of more popular interfaces like USB and RS-232 [18].

The IEEE 1451.3 standard establishes a distributed multidrop interface network sharing a common pair of wires. This standard specifies multiple TIMs named TBIM or Transducer Bus Interface Modules, and it is intended to allow synchronized reading of large sensor arrays on a parallel transducer bus. Multidrop connectivity is used by defining channel identification protocols, hot-swap protocols, time synchronization protocols, and the read and write logic functions used to access the TEDS and transducer data. The standard is now being revised to be adapted to the requirements of the IEEE 1451.0 standard.

The IEEE 1451.4 standard presents a mixed-mode interface (MMI) for analog transducers. TEDS to be used with this standard are also defined due to their specificity as they do not comply with TEDS defined in the IEEE 1451.0 standard [19]. The IEEE 1451.4 standard defines a mechanism for adding self-identification technology to traditional analog sensors and actuators.

The IEEE 1451.5 standard defines wireless communication methods between TIM and NCAP. Standards such as 802.11 (WiFi), 802.15.1 (Bluetooth), 802.15.4 (ZigBee), and 6LowPAN are adopted as viable wireless interface. Finally, other well-known interfaces, the CANopen interface and the Radio Frequency Identification (RFID) interface, meet, or will meet soon, the standard.

The proposed IEEE P1451.6 standard establishes a transducer-to-NCAP interface using the CANopen network. It also defines a mapping of the IEEE 1451 TEDS to the CANopen dictionary entries.

The recently approved IEEE 1451.7 standard describes communication methods and data formats, providing a TEDS for RFID sensors.

The TEDS is a key concept of the IEEE 1451 standard [20]. It standardizes transducers into TIMs, so any application can dynamically discover equipments, accessing their data and characteristics. TEDSs are classified into mandatory and optional TEDSs, following the IEEE 1451.0 specifications. The mandatory TEDSs include the Meta TEDSs, the Transducer Channel TEDSs, the PHY TEDSs, and the User Transducer Name TEDSs. In addition to these essential TEDSs, the IEEE 1451.0 standard supports other TEDSs: Calibration TEDS, Frequency Response TEDS, Transfer Function TEDS, Manufacturer-Defined TEDS, End User Application-specific TEDS, and Text-Based TEDS (including Meta ID TEDS, Transducer Channel ID TEDS, Calibration ID TEDS, Command TEDS, Location and Title TEDS, and Geolocation TEDS).

The IEEE 1451.0 standard specifies the format for these TEDSs, as it is shown in Table 1. This format is flexible and

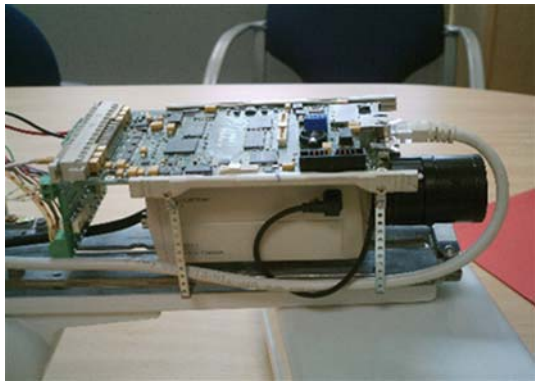


Figure 3: VisioWay prototype.

Figure 4: Detection areas in an installed prototype: details of the directional regions behavior (R3, R4, and R5) and queue detection areas (R1 and R2).

extensible to handle a wide variety of sensors, and it is compact to save memory space in the system. Each TEDS has a 4-octet length field, a data block with a variable size depending on the TEDS, and a 2-octet checksum field to verify the integrity of the TEDS. The data block includes different entries, and it uses a Type/Length/Value (TLV) structure to store each entry. The "Type" field is a 1-octet tag that identifies the data, the "Length" field specifies the number of octets of the data, the "Value" field includes the actual data.

The information in the data block depends on the TEDS, except in the case of the "TEDSID" field or TEDS identifier header. This field is standard in all TEDSs, including properties such as family (always 0), class (TEDS access code), version (IEEE 1451.0 standard version), and the number of octets in the length field of all TLV tuples in the TEDS except this one (normally 1).

The TEDS is normally stored in nonvolatile memories in the TIM, while the NCAP has a transferred copy. If the TEDS is not stored in the TIM, it is called virtual TEDS and the manufactures are expected to provide them. The standard allows the users to program some of these TEDSs once the smart sensor is installed, anyway. Hand assembly of TEDS is a very tedious task, even for the most simple TIMs. Hence, a TEDS compiler is highly recommended for this purpose.

Many smart sensor applications using IEEE 1451 are being developed [21, 22], including web-based sensor systems [23–26]. In this paper, the utility of the IEEE 1451 stand-ard for the development of ubiquitous electronic equipment networks in road-traffic applications is analyzed. In particular, an existing video-based traffic sensor, VisioWay [27], is taken as a starting point to build an IEEE-1451 compliant smart traffic sensor, capable of providing new web services to the traffic control centers.

3. VisioWay in the Road-Traffic Management Field

VisioWay is a prototype of a novel electronic equipment for application in the ITS field using artificial vision [27]. Figure 3 shows the equipment. It has been designed for commercial purposes in traffic infrastructures, providing web services to the users of the infrastructure and preserving its robustness as smart sensor in IP-based networks. VisioWay was designed as a road-traffic parameter estimation system, providing useful traffic information such as traffic flow, lane average speed and occupancy, or congestion levels. The data collection is structured according to a number of user-configurable polygons with arbitrary shape or size, called detection areas (Figure 4). Each one of these detection areas has interesting functionalities like presence and directional detection functionalities. The automatic vehicle detection method implemented in VisioWay is based on advanced background subtraction methods [28, 29].

VisioWay has the following features.

(i) The Freescale i.MX21 multimedia processor is the core of the hardware system.

(ii) Wire (Ethernet) and wireless (Bluetooth) connectivity is included.

(iii) Interface to typical data storage media like USB or SD/MMC is provided.

(iv) The hardware is driven by embedded Linux OS (kernel 2.4.20).

(v) Artificial vision algorithms have been implemented for gathering information about relevant foreground objects (typically vehicles) in a traffic scene.

The i.MX21 processor is based on the advanced and power-efficient RISC processor ARM926EJ-S core, operating at speeds up to 266 MHz. On-chip modules are provided, including LCD and MMC/SD controllers, USB controllers, CMOS sensor interface, and an enhanced MultiMedia Accelerator (eMMA). The CMOS sensor interface provides the capability to acquire digital images, typically BT656 or raw data streams in the RGB or YUV components, delivering them to the media processor. The eMMA module consists of video processor units and encoder/decoder modules that support MPEG-4 and H.263 real-time encoding/decoding of images from 32×32 pixels up to CIF format at 30 fps. The processing unit is based on a preprocessor module that resizes input frames from memory or from the CMOS sensor

Figure 5: VisioWay embedded system.

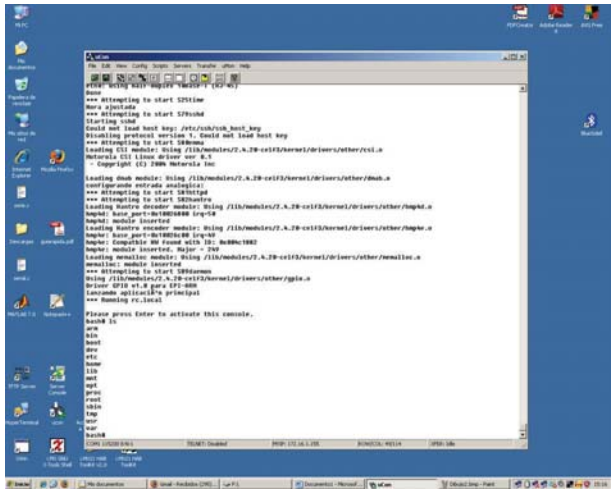


Figure 6: Kernel prompt of the prototype.

interface, performing color space conversion, and a postprocessor module that takes raw images from memory performing additional processing of a MPEG-4 video streaming to deblock, dering, resize, or color space conversion on decoded frames.

The prototype board (Figure 5) is also provided with a mini USB-OTG connector, a SD card connector, a RJ45 Ethernet connector, a Bluetooth expansion connector, and an analog video connector owed by a video decoder chip to allow analog video (PAL and NTSC formats) processing. Onboard memory allows video processing and OS storage. A 32 M $\times$ 16 Flash memory is used to store the OS kernel image while two 16 M $\times$ 16 SDRAM modules are used for video processing and for decompressing the kernel image. An ARM Linux (kernel 2.4.20) operates the i.MX21 video processor, allowing access to a wide variety of open source software modules (Figure 6).

Several processes and services are in concurrent operation with the main process (road-traffic monitoring application). These processes are a watchdog process for rebooting the system when a periodic signal is not received from the main process, a MPEG server that operates as an interface to the i.MX21 MPEG hardware module for frame feeding from the application, and a HTTP server for configuration and supervision purposes, including an SSH server for remote logging into the system, and an FTP server for upload and download operations.

4. VisioWay Smart Sensor and Experimental Results

The VisioWay system must be adapted to comply with the IEEE 1451 standard. The original HTTP and FTP services, adopted for reading data and configuring the VisioWay equipment, cannot provide by themselves a simple way to access all the system information, especially when plug-and-play capabilities are expected. This problem affects the development of applications that makes use of this data in two different ways. First, the application is dependent on the data structures and presentation forms and thus needs to be modified according to future changes. Second, it makes the development of new applications more difficult since a deeper knowledge of the device is required. As the devices are intended to estimate and store traffic measurements, it seems reasonable to consider them as multichannel sensors, interconnected into a sensor network. VisioWay was also intended to have different functionalities and to be installed in different locations over the urban area. Consequently, it is essential to have a common data structure and interface in order to guarantee full compatibility between different VisioWay equipments, making necessary a unique database with all the information that should be provided to upper-level systems and applications.

A network-compatible smart sensor node is implemented in VisioWay using the IEEE 1451 family of standards. The objective is to group VisioWay equipments into a single clustered network, some of them behaving as cluster heads—NCAPs in the IEEE 1451 notation. A common database (TEDS) must be designed, allowing connection regardless of the interface type. The intelligence of the device obviates unnecessary tasks when making the connection between systems, providing the interface information at connexion time, as a plug- and play-function. Other implementation details are as follows.

(1) The system has been designed according to the IEEE 1451.0 standard specifications since this standard incorporates all the specifications except for IEEE 1451.4.

(2) VisioWay can implement different functionalities, described in each Transducer Channel TEDS.

(3) The role perspective: since VisioWay is intended to be installed across the city and connected though TCP/IP networks, each device can be considered as the NCAP of a particular sensor network using the HTTP API defined in the IEEE 1451.0 standard. At the same time, it is possible to connect other VisioWay devices into the network as TIMs, using the IEEE 1451.x standards and providing the NCAP a gateway function between TIMs and the Internet. The following alternatives are possible according to the developed implementation.

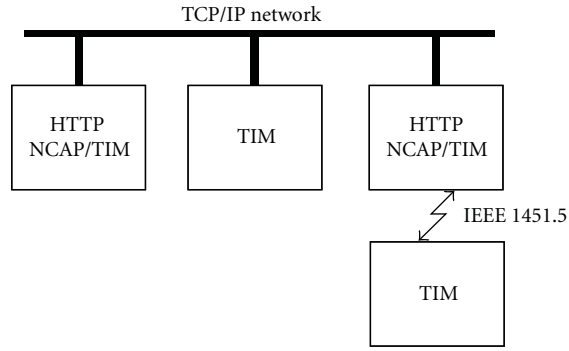


Figure 7: VisioWay equipments into the same TCP/IP network, playing different roles (NCAP and/or TIM roles).

(a) A VisioWay system integrates a NCAP and a TIM into a single device. The Inter-Process Communication mechanisms available in the embedded Linux operating system link the NCAP and the TIM. As explained before, the VisioWay system is running under an ARM Linux kernel [27].

(b) A VisioWay system serves as NCAP for other devices (TIMs using IEEE 1451.x interfaces). The current implementation supports the Bluetooth standard, as specified in the IEEE 1451.5 standard.

(c) TCP/IP-based TIMs can be connected into the TCP/IP network. This approach has the limitation of being only standardized for the IEEE 802.11 interfaces via the IEEE 1451.5 standard. Nevertheless, this option makes sense since the communications are based on the IEEE 1451.0 message format, and the aim of the IEEE 1451.0 standard is to encourage compatibility across the IEEE 1451 family, making easier the addition of new interfaces, especially when TCP/IP is used.

According to the previous paragraphs, different VisioWay systems, playing different roles, can be connected into the same TCP/IP network, as it is shown in Figure 7. Upper-level applications need only to distinguish between NCAP and TIM access. Consequently, if a final user application is communicating with a TIM or an NCAP, IEEE 1451.0 or HTTP messages are, respectively, exchanged. Notice that every NCAP can be connected to multiple TIMs.

4.1. NCAP Implementation. An IEEE 1451.0 HTTP server has been developed in C language to implement NCAP functions over the VisioWay Linux kernel. Figure 8 depicts the block diagram of the designed NCAP application, while Table 2 shows all the implemented commands.

Although VisioWay, in its original form, runs a general purpose HTTP server, it is preferred to develop a new IEEE 1451.0 compliant HTTP server, taking into account the communication with multiple TIMs over different physical interfaces and maintaining a minimum CPU and memory usage. The information transfer (HTTP and IEEE 1451.0) is managed using the Linux select system call, avoiding the need for multiple threads. Notice that more complex configurations are allowed running several HTTP servers within the same process. When the implemented HTTP server receives a request, it is analyzed to check if it is an implemented command. IEEE 1451.0 commands are dispatched to TIMs if required and TIMs are controlled to perform measurements or accessed from the client to read the results in the IEEE 1451.0 format. Notice that each registered TIM has a queue of requests in the NCAP implementation, and that each TIM request can be replied or not. If a reply is not expected, the NCAP uses the timing parameters included in the Meta-TEDS of the TIM to determine whether the TIM is ready to receive new commands. In the case of being ready, the reply is retrieved from the queue of requests. On the contrary, if a reply is expected, the next command is sent only after its reception, taking into account timeout errors to consider failed operations that are reported to the HTTP client. It is also possible to prioritize commands, for example, reading a sensor against writing a TEDS, which requires more processing time.

4.2. TIM Design. The TIM implementation is based in the IEEE 1451.0 standard, allowing different physical interfaces. Eight standard command classes (0–7), some reserved classes (8–127), and open classes to manufacturers (128–255) are included, and the TIM is implemented as an independent process with the ability to manage and store the basic information and functionality of the sensor device using TEDS forms. After initialization, the TIM runs like a data source server. It checks the physical interface for different requests from the NCAP and decodes the command, which indicates the type of function requested and the channel address. The most common functions include reading of the channel transducer data, reading data types, or reading the channel status. If an error is detected on any of the Transducer Channels during the service, the TIM sets an appropriate flag in the channel status register.

All messages between TIM and NCAP can be classified into command messages, TIM initiated messages and reply messages, following the IEEE 1451.0 standard recommendations. All transmitted information is bundled together into a payload, which is then encoded into an octet array.

The IEEE 1451.0 and 1451.5 standards allow the implementation of a point-to-point (P2P) or general networked connection for the TIM. A P2P connection is chosen because each TIM does not normally talk with more than one NCAP, allowing for a simpler TIM implementation, as neither addressing nor discovery protocols are required.

4.3. TEDS Realization. A simple TEDS compiler and a TEDS reader have been developed to generate and test the TEDS files for the VisioWay equipment. The TEDS compiler appli-cation is a WEB-based interface that allows users to introduce the required parameter fields and generates the TEDS accordingly. The generated TEDS is displayed as

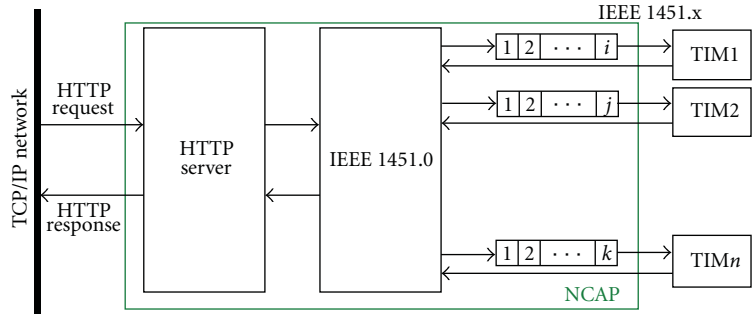


Figure 8: Block diagram of the implemented NCAP application (IEEE 1451.0-compatible HTTP server).

Table 2: Implemented commands.

Module	Operation	Task
Discovery API	TimDiscovery	List available TIMs
	TransducerDiscovery	List channels in one TIM
Transducer Access API	ReadData	Read access to measured data in a specific channel
	WriteData	Write access to a specific channel
TEDS Manager API	ReadTeds	Read a specific TEDS (Access Code)
	ReadRawTeds	Read TEDS from TIM bypassing NCAP cache
	WriteTeds	Write a specific TEDS
	WriteRawTeds	Write TEDS in TIM bypassing NCAP cache
	UpdateTedsCache	Update TEDS Cache
Transducer Manager API	Trigger	Perform a trigger on the specific channel

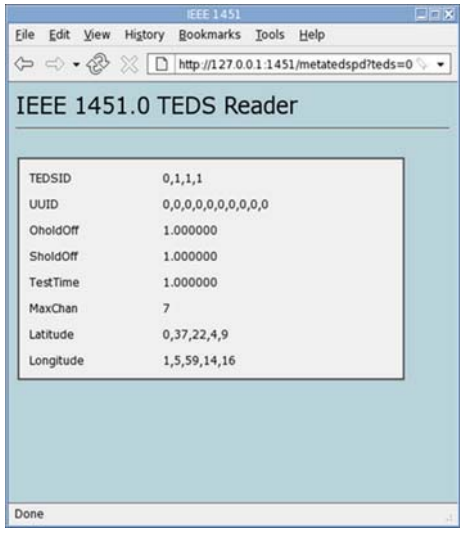


Figure 9: TEDS reader application.

a comma-separated list of bytes, which can be used to initialize memory buffers in the TIM program source code. The TEDS reader application is another WEB-based interface, as it is shown in Figure 9.

The implemented Meta TEDS describes the relations among the different Transducer Channels of the TIM and stores the timing parameters in the worst-case scenario. It indicates the number of implemented Transducer Channels, being possible to group them into control, vector, geographic location, and transducer channel proxies groups. The timing parameters are used by the NCAP to establish time-out values in the communication software and determine a nonresponse event from the TIM. This TEDS also includes an "UUID" field, or globally unique identifier, which is an identification data associated with the TIM whose value is unique and it can include some manufacturing details.

Some fields are available for the manufacturers. Two of these fields have been used to save the geostationary position of the device: latitude (type 150) and longitude (type 151). Each field includes 5 octets, the first one being used to indicate the sign (i.e., north/south or east/west), while the rest specifies the degrees, minutes, seconds, and hundredths of second of the position. This format has been adopted for simplicity reasons.

Figure 10 shows the TEDS compiler application defining the Meta TEDS. The obtained file in HEX format is shown in Figure 11, including the total length in octets, the checksum, and the TLV tuples of the data block. At this stage, the Meta TEDS must be stored in the memory system of the TIM.

The *Transducer Channel TEDS* shows all the characteristics of the available sensors (each sensor has an associated Transducer Channel TEDS). VisioWay has been adapted in such a way that each functionality (vehicle-presence detection, directional vehicle detection, vehicle-queue detection, vehicle counting, occupancy, and traffic incident detection) is considered as an independent sensor, having their own

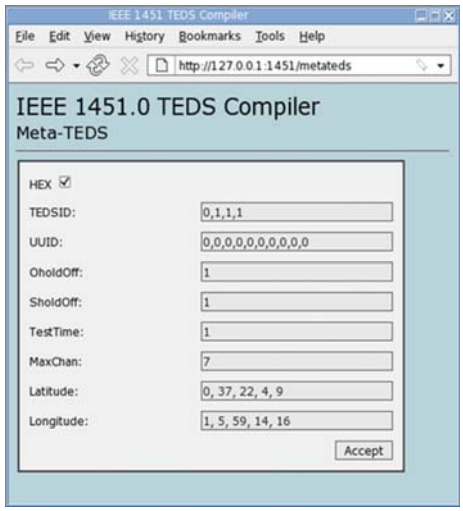


Figure 10: TEDS compiler application generating the Meta TEDS.

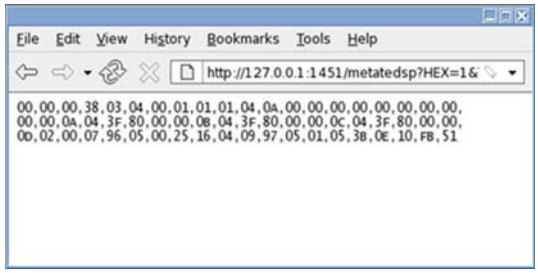


Figure 11: Meta TEDS "file" in HEX format.

Transducer Channel TEDS. The following information is stored in the data block of the TEDS.

(i) The nature of the measured data (e.g., physical units and lower and upper range limits).

(ii) The format of the data to be read from the TIM. This includes the data type (e.g., UInt8 or Float32) or the number of individual data samples in each data set, specified in fields like data model, data model length, model significant bits, and maximum data repetitions.

(iii) The timing parameters describing the maximum required time to perform different procedures (e.g., transducer channel read setup time or self-test time), as well as the sensor sampling period.

(iv) The available sampling modes of the sensor. Different sampling modes are supported in the standard (continuous, free-running, immediate, and trigger-initiated), as well as multiple buffers. Relevant types (i.e., fields) for these properties are sampling attribute, sampling mode capability, default sampling mode, buffered attribute, and data transmission attribute.

Figure 12 shows the TEDS compiler application creating the Transducer Channel TEDS. This TEDS is flexible enough

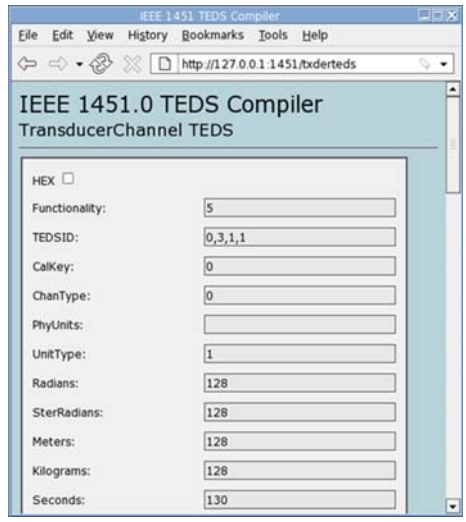


Figure 12: TEDS compiler application generating the Transducer Channel TEDS.

to accommodate to different transducers. The standard defines a small set of basic units that can be combined to provide most of the existing physical units. This set includes the radian and steradian and the seven international system base units. Any particular unit used in a smart sensor can be defined as a mathematical combination of these set of units. In particular, occupancy estimation, percentage of full queue, and other dimensionless quantities are obtained using VisioWay; thereby, the base units field is set to 0. Other measurements can be obtained from the ratio of two quantities, and their units can be specified using the Physical Units Interpretation field of this TEDS. Float32 data types have been chosen to be used in VisioWay channels.

Other characteristics and specifications defined with this TEDS for the smart sensor based on VisioWay are the following

(i) Free running with pretrigger as sampling mode, although the immediate mode is also possible.

(ii) One buffer is used.

(iii) The system returns data when instructed. The smart sensor does not contain any streaming capability, and the data transmission attribute in this TEDS, which establishes the available transmission modes (streaming mode or request-triggered mode), is omitted. Nevertheless, streaming support could be included in future developments.

(iv) The equipment boot time was used as the warm-up time. Notice that timing parameters are highly dependent on the hardware and optimizations of the VisioWay software. For example, fields like Transducer Channel update time, which indicates the maximum time between a trigger or a read command reception and the data available event in the data set, are determined by the video processing algorithms. Other parameters like Transducer Channel read delay time, which indicates the maximum delay between

Table 3: New types for the programmed Transducer Channel TEDS of VisioWay.

Functionality	Value
Vehicle-presence detection	1
Vehicle directional detection	2
Traffic incident detection	3
Vehicle counting	4
Occupancy	5

a read command reception and the corresponding start-transmission event, are also difficult to evaluate in the VisioWay equipment.

The types that the standard defines for this TEDS do not allow a complete description of VisioWay. The plug-and-play mechanism of VisioWay requires the use of one of the open-to-manufacturer types in the Transducer Channel TEDS (Table 3). Notice that more ITS-based services can be used, increasing the utility of the proposed system. For instance, if the state of a traffic light is available in real time to the VisioWay equipment, applications like red-light runner detection are feasible.

The *PHY TEDS* is dependent on the physical communication media that connects the TIM and the NCAP although the data are read-only octets after the TED is programmed. It includes detailed information of the IEEE 1451.X physical layer, establishing the minimum transmit latency, the maximum connected devices, and the maximum data throughput characteristics in the case of a wireless link. New types were defined to identify the IP and the physical addresses of the VisioWay wireless module. Notice that the information contained in this TEDS is not essential to access the TIM since most physical layers provide auto-configuration degrees completely independent of the IEEE 1451 standard.

The *User Transducer Name TEDS* stores the TIM name (the name assigned by an end-user to a specific TIM), establishing the name to identify a specific transducer in the complete system. The structure of this TEDS is recommended in the standard, and the manufacturer must provide a blank space of nonvolatile memory that can be programmed by the end-user using the standard TEDS access methods.

4.4. Traffic Monitoring Application Based on VisioWay Smart Sensor Equipments. The IEEE 1451 family of standards allows using VisioWay for providing real-time traffic data access to be used in web-based traffic control and information applications. Different prototypes are installed across the city of Seville and connected to a private Local Area Network (LAN) using its 10/100 Mbps Ethernet port and a RJ45 connector. Figure 13 shows two equipments monitoring a crossroads. The link speed and the device throughput on the private LAN have been analyzed.

A thorough study of the VisioWay capabilities and performance is required, prior to its conversion into a smart transducer. A multicast control mechanism must be applied to make an effective use of the available bandwidth [30]. An

Figure 13: VisioWay in two real emplacements in the city of Seville, Spain.

experiment has been performed to prove the robustness of the prototype in congested networks with multicast video delivery. One external PC runs the network traffic monitoring software, called monitoring program, based on libpcap libraries (Tcpdump/libpcap project). Libpcap utility is widely used for low-level packet monitoring and filtering. The main idea of this utility is to capture SOAP incoming and outgoing packets, while being completely independent from client coding. Using libpcap utility, each bit sent or received through the physical link can be captured, allowing the monitoring of the whole LAN traffic. Other set of PCs work as Data Servers. Two tests have been performed.

(i) The congested network is emulated using the data servers, and the performance of VisioWay is studied. First, the throughput of the main algorithm—processed frames per second (FPS) when performing typical vehicle detection tasks—is evaluated under congested networks (up to 8000000 bps of network traffic) with a multicast control mechanism. The results show that the system throughput does not degrade in these conditions, remaining at steady values around 7 FPS. Notice that this ability depends on the hardware design. The Ethernet controller included in VisioWay analyzes network traffic at the network layer, avoiding any overload of the VisioWay CPU.
The performance of VisioWay has also been tested under more adverse conditions, such as unicast control mechanism, with a 8000000 bps network traffic addressed to VisioWay. In this case, the generated network traffic degrades the system throughput till 5,57 FPS.

(ii) Second, VisioWay performance is analyzed under constant real-time video delivery, using UDP protocol. The real-time video (MPEG-4 elementary video using CIF format, at 384 Kbps and 25 FPS) is delivered to the network monitoring PC. Table 4 and Figure 14 summarize the obtained results. The time between packets is acceptable for most MPEG-4 players in the 85% cases without necessity of incoming buffers. These buffers, usually included in

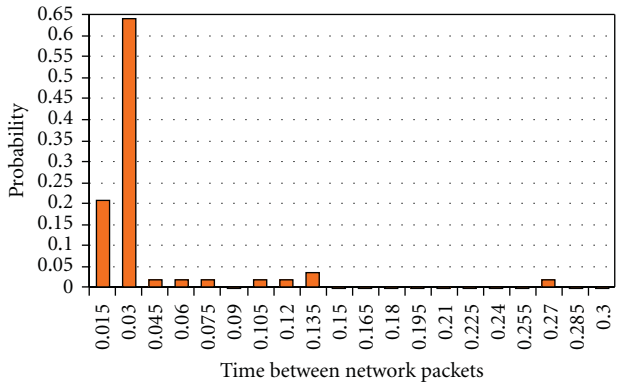


Figure 14: VisioWay without the smart sensor capability. Network traffic associated with video packets: time between packets probability density function.

Table 4: Performance of VisioWay without the smart sensor capability as a real-time video supplier.

Variables	Value
No. of packets	53
Average	0,030153 s
Standard deviation	0,043664 s
Maximum	0,260425 s
Minimum	0,000080 s

conventional MPEG-4 players, relax the random restrictions associated with the time between network packets.

As a conclusion, the device throughput on the Internet is robust. As expected, heavy network traffic is obtained when real-time video streaming is demanded. Figure 15 details the running processes on the VisioWay and their CPU time and memory requirements. The running processes correspond to the video processing algorithms (epi.out) and the SSH server for remote logging. In particular, the first one consumes the majority of the CPU time (98,8%), but just a small amount of memory (4,1%).

The developed IEEE 1451 HTTP interface allows an easy and straightforward way to acquire real-time traffic data, abstracting the implementation details and simplifying the development of web-based applications. A small web application has been developed to analyze the viability of introducing HTTP-NCAP services for monitoring road traffic in the city of Seville, Spain. The implemented application represents into Google Maps one of the specific traffic-data parameters obtained in VisioWay equipment related with traffic conditions: the region occupancy. This information is stored into designed Meta-TEDS. The application has been developed using Javascript, while the web access to the IEEE 1451.0 HTTP API is achieved using the XMLHttpRequest object. In this way, no server-side processing is required and the script supports the computational load of the response. Figure 16 details the VisioWay equipments used in the experiment, showing also the obtained occupancy values and

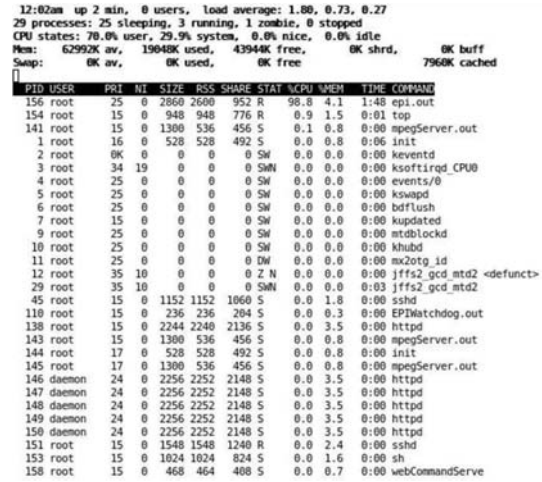


Figure 15: CPU time and memory requirements of concurrent processes executed in the VisioWay equipment prior its conversion into a smart transducer equipment.

Figure 16: Overview of the web application for identifying occupancy data around the city of Seville (Spain). Three VisioWay smart transducer equipments (TIM ID from 1 to 3) are shown on a Google Map of the analyzed area.

the Meta-TEDS of two of the equipments (TIM ID 3 and 1, resp.). Figure 17 shows the VisioWay equipments playing the "vehicle counting" task.

The device throughput on the Internet is more robust than before. The implementation of the NCAP and TIM processes corresponds to a 0% of the CPU time, being the CPU time requirements similar to the previous case (VisioWay equipment prior its conversion into a smart transducer) (Figure 15). Notice that the required amount of memory for the NCAP and TIM implementations is about 208 kbyte and 72 kbyte, respectively, including shared libraries (less than 3,17% and 1,10% of the total available memory, and about 25% and 15% of the video processing algorithms, resp.).

5. Conclusions

The development of the IEEE 1451 standard offers new opportunities for the ITS technology and embedded systems

Figure 17: One VisioWay smart transducer implementing its vehicle-counting functionality, integrated into Google Map environment.

with application in road traffic management. The IEEE 1451 family of standards has been used to provide an open platform for the development of networked equipments, simplifying the connectivity to a network over the basis of HTTP web services and providing "plug-and-play" capabilities. In this paper, the combined use of the IEEE 1451 standard, the Internet, and embedded equipments has been analyzed in the context of web-based road-traffic monitoring applications. The IEEE 1451 standard has been integrated into an automatic video processing system, commercialized under the name of VisioWay. The proposed architecture has been tested on an experimental setup with the purpose of monitoring road-traffic data in the city of Seville, Spain, using the Internet as the communication link. The obtained results prove the effectiveness of the proposed system in the development of future ITS equipments.

Acknowledgments

VisioWay is a trademark of ACISA. This work has been supported by the Spanish Firm ACISA, the Spanish Ministry of Education and Science (Research Project with reference DPI2009-07955), and the Consejería de Innovación, Ciencia y Empresa (Research Project with reference P07-TIC-02621).

References

[1] C. Y. Chong and S. P. Kumar, "Sensor networks: evolution, opportunities, and challenges," *Proceedings of the IEEE*, vol. 91, no. 8, pp. 1247–1256, 2003.

[2] R. Aguilar-Ponce, A. Kumar, J. L. Tecpanecatl-Xihuitl, and M. Bayoumi, "A network of sensor-based framework for automated visual surveillance," *Journal of Network and Computer Applications*, vol. 30, no. 3, pp. 1244–1271, 2007.

[3] 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.

[4] M. Can Filibeli, O. Ozkasap, and M. Reha Civanlar, "Embedded web server-based home appliance networks," *Journal of Network and Computer Applications*, vol. 30, no. 2, pp. 499–514, 2007.

[5] P. Kulkarni and Y. Ozturk, "MPHASiS: mobile patient healthcare and sensor information system," *Journal of Network and Computer Applications*, vol. 34, no. 1, pp. 402–417, 2011.

[6] S. L. Toral, M. R. M. Torres, F. J. Barrero, and M. R. Arahal, "Current paradigms in intelligent transportation systems," *IET Intelligent Transport Systems*, vol. 4, no. 3, pp. 201–211, 2010.

[7] F. Barrero, S. Toral, M. Vargas, F. Cortés, and J. M. Milla, "Internet in the development of future road-traffic control systems," *Internet Research*, vol. 20, no. 2, pp. 154–168, 2010.

[8] P. Alexander, D. Haley, and A. Grant, "Cooperative intelligent transport systems: 5.9-GHz field trials," *Proceedings of the IEEE*, vol. 99, no. 7, pp. 1213–1235, 2011.

[9] L. Chen, C. Jiang, and J. Li, "VGITS: ITS based on intervehicle communication networks and grid technology," *Journal of Network and Computer Applications*, vol. 31, no. 3, pp. 285–302, 2008.

[10] M. Esteve, C. E. Palau, J. Martínez-Nohales, and B. Molina, "A video streaming application for urban traffic management," *Journal of Network and Computer Applications*, vol. 30, no. 2, pp. 479–498, 2007.

[11] M. Vargas, J. M. Milla, S. L. Toral, and F. Barrero, "An enhanced background estimation algorithm for vehicle detection in urban traffic scenes," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 8, pp. 3694–3709, 2010.

[12] R. Verdone, D. Dardari, G. Mazzini, and A. Conti, *Wireless Sensor and Actuator Networks: Technologies, Analysis and Design*, Academic Press, 2007.

[13] E. Y. Song and K. Lee, "Understanding IEEE 1451—networked smart transducer interface standard—what is a smart transducer?" *IEEE Instrumentation and Measurement Magazine*, vol. 11, no. 2, pp. 11–17, 2008.

[14] V. Viegas, J. M. Dias Pereira, and P. M. B. Silva Girao, "A brief tutorial on the IEEE 1451.1 standard—part 13 in a series of tutorials in instrumentation and measurement," *IEEE Instrumentation and Measurement Magazine*, vol. 11, no. 2, pp. 38–46, 2008.

[15] E. Y. Song and K. B. Lee, "Smart transducer web services based on the IEEE 1451.0 standard," in *Proceedings of the IEEE Instrumentation and Measurement Technology (IMTC '07)*, May 2007.

[16] E. Y. Song and K. B. Lee, "STWS: a unified web service for IEEE 1451 smart transducers," *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 8, pp. 1749–1756, 2008.

[17] Y. Wang, M. Nishikawa, R. Maeda, M. Fukunaga, and K. Watanabe, "A smart thermal environment monitor based on IEEE 1451.2 standard for global networking," *IEEE Transactions on Instrumentation and Measurement*, vol. 54, no. 3, pp. 1321–1326, 2005.

[18] E. Y. Song and K. B. Lee, "Sensor network based on IEEE 1451.0 and IEEE p1451.2-RS232," in *Proceedings of the IEEE International Instrumentation and Measurement Technology Conference (I2MTC '08)*, pp. 1728–1733, May 2008.

[19] N. Ulivieri, C. Distante, T. Luca, S. Rocchi, and P. Siciliano, "IEEE1451.4: a way to standardize gas sensor," *Sensors and Actuators B*, vol. 114, no. 1, pp. 141–151, 2006.

[20] S. Manda and D. Gurkan, "IEEE 1451.0 compatible TEDS creation using.NET Framework," in *Proceedings of the IEEE Sensors Applications Symposium (SAS '09)*, pp. 281–286, February 2009.

[21] L. Bissi, P. Placidi, A. Scorzoni, I. Elmi, and S. Zampolli, "Environmental monitoring system compliant with the IEEE 1451 standard and featuring a simplified transducer interface," *Sensors and Actuators A*, vol. 137, no. 1, pp. 175–184, 2007.

[22] D. Wobschall and S. Mupparaju, "Low-power wireless sensor with SNAP and IEEE 1451 protocol," in *Proceedings of the 3rd IEEE Sensors Applications Symposium (SAS '08)*, pp. 225–227, February 2008.

[23] A. Flammini, P. Ferrari, E. Sisinni, D. Marioli, and A. Taroni, "Sensor integration in industrial environment: from fieldbus to web sensors," *Computer Standards and Interfaces*, vol. 25, no. 2, pp. 183–194, 2003.

[24] E. F. Sadok and R. Liscano, "A web-services framework for 1451 sensor networks," in *Proceedings of the IEEE Instrumentation and Measurement Technology Conference (IMTC '05)*, pp. 554–559, May 2005.

[25] G. Bucci, F. Ciancetta, E. Fiorucci, D. Gallo, and C. Landi, "A low cost embedded web services for measurements on power system," in *Proceedings of the IEEE International Conference onVirtual Environments, Human-Computer Interfaces, and Measurement Systems (VECIMS '05)*, pp. 7–12, June 2005.

[26] K. B. Lee and E. Y. Song, "Sensor Alert Web Service for IEEE 1451-based sensor networks," in *Proceedings of the IEEE Intrumentation and Measurement Technology Conference (I2MTC '09)*, pp. 1055–1058, May 2009.

[27] S. L. Toral, M. Vargas, and F. Barrero, "Embedded multimedia processors for road-traffic parameter estimation," *Computer*, vol. 42, no. 12, pp. 61–68, 2009.

[28] S. Toral, M. Vargas, F. Barrero, and M. G. Ortega, "Improved sigma-delta background estimation for vehicle detection," *Electronics Letters*, vol. 45, no. 1, pp. 32–34, 2009.

[29] J. M. Milla, S. L. Toral, M. Vargas, and F. Barrero, "Computer vision techniques for background modelling in urban traffic monitoring," in *Urban Transport and Hybrid Vehicles*, S. Sciyo, Ed., 2010.

[30] W. Q. Sun, J. S. Li, and P. L. Hong, "A stateful multicast access control mechanism for future metro-area-networks," *Internet Research*, vol. 13, no. 2, pp. 134–138, 2003.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 234029, 7 pages
doi:10.1155/2012/234029

Research Article

Traffic Rerouting Strategy against Jamming Attacks in WSNs for Microgrid

Yujin Lim,[1] Hak-Man Kim,[2] and Tetsuo Kinoshita[3]

[1] *Department of Information Media, University of Suwon, San 2-2 Wau-ri, Bongdam-eup, Hwaseong-si, Gyeonggi-do 445-743, Republic of Korea*
[2] *Department of Electrical Engineering, University of Incheon, 12-1, Songdo-dong, Yeonsu-gu, Incheon 406-840, Republic of Korea*
[3] *Department of Computer and Mathematical Sciences, Graduate School of Information Science, Tohoku University, Sendai 980-8577, Japan*

Correspondence should be addressed to Hak-Man Kim, hmkim@incheon.ac.kr

Received 4 November 2011; Accepted 25 February 2012

Academic Editor: Carlos Ramos

In a microgrid as an energy infrastructure, the vulnerability against jamming attacks is fatal. Thus, the ability to deal with jamming attacks and maintain an acceptable level of service degradation in presence of the attacks is needed. To solve the problem, we propose a traffic rerouting scheme in wireless communication infrastructure for islanded microgrid. We determine disjoint multiple paths as candidates of a detour path and then select the detour path among the candidates in order to reduce the effect of jamming attack and distribute traffic flows on different detour paths. Through performance comparison, we show that our scheme outperforms a conventional scheme in terms of packet delivery ratio and end-to-end delay.

1. Introduction

A microgrid is a localized grouping of electricity generation, energy storage, and loads, and it is normally connected to an upstream power grid [1–4]. However, by occurrence of fault occurrence in the power grid or by geographical isolation, a microgrid can be isolated from the power grid, and it is called an islanded microgrid [5]. Main goal of a microgrid operation is to balance the power between power supply and power demand. The severe power imbalance may happen due to power system faults and protection system malfunctions, and it may eventually lead to system collapse through a frequency instability process. Literatures [6–12] have employed a multiagent system to operate a microgrid economically and efficiently. A multiagent system is a system composed of multiple interacting intelligent agents to solve problems that are difficult or impossible for an individual agent or a monolithic system to solve [13]. It is a good solution for autonomous operation of a microgrid composed of distributed devices and systems. Figure 1 shows the islanded microgrid based on the multi-agent system. Each component has an agent, and the agent exchanges information among other agents. Especially, the agent in the microgrid operation & control center (MGOCC) gathers information about the supplied power and the load demands and operates the microgrid to balance the power.

In this paper, we consider a communication infrastructure based on the wireless sensor network (WSN) for geographically islanded microgrid operated and controlled by the multi-agent system. As an extension of the WSN, we employ a wireless mesh network (WMN) [14]. The WMN has been recently developed to provide high-quality services and applications over wireless personal area networks, wireless local area networks, and wireless metropolitan area networks [15]. The WMN has a hybrid network infrastructure with a backbone and an access network. It is operated in both ad hoc and infrastructure modes with self-configuration and self-organization capabilities. The WMN has been envisioned as the economically viable networking paradigm to build up broadband and large-scale wireless commodity networks. Installing a cabling infrastructure not only slows down implementation but also significantly increases installation cost. On the other hand, building a WMN enormously reduces the infrastructural cost because the mesh network

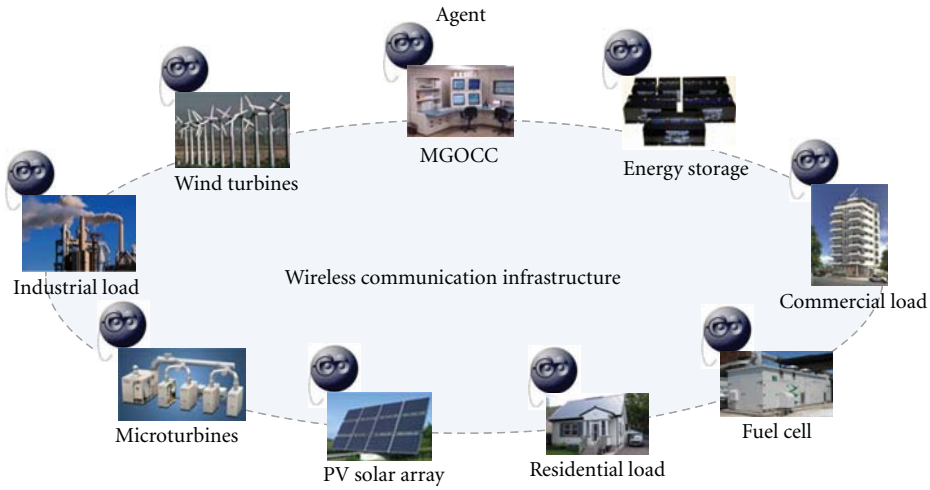


Figure 1: The islanded microgrid based on multiagent system.

needs only a few access points for connection. This reduction of network installation cost ensures rapid deployment of a metropolitan broadband network even in rural or scarcely populated urban areas. Thus, we employ the WMN to design a communication infrastructure for the islanded microgrid.

However, built upon open wireless medium, the WMN is particularly vulnerable to jamming attacks [16]. In the microgrid as an energy infrastructure, the vulnerability is a critical problem. Thus, the ability to deal with jamming attacks and to maintain an acceptable level of service degradation in presence of jamming attacks is a crucial issue in the design of the WMN. To solve the problem, traffic rerouting, channel reassignment, and scheduling schemes have been considered as jamming defense strategies. In this paper, we propose a traffic rerouting scheme in the WMN. To reduce the effect of jamming attacks, we determine multiple candidates of a detour path, and the multiple candidates are physically disjoint. Once the candidates are given, to distribute traffic flows on different detour paths, we stochastically select one candidate path as a detour path. We verify that our scheme improves the packet delivery ratio and end-to-end delay in comparison with a conventional scheme.

The remainder of this paper is structured as follows. Section 2 introduces a jamming attack and defines our WMN architecture. Section 3 explains our traffic rerouting scheme. Following this, we verify the proposed scheme by using NS-2 simulator in Section 4. Finally, Section 5 summarizes our study results.

2. System Model

2.1. Jamming Attack. Jamming represents the most serious security threat in the field of a wireless communication. Jamming is defined as the act of intentionally directing electromagnetic energy towards a communication system to disrupt or prevent signal transmission [17]. In the context of a wireless communication, jamming is the type of attack which interferes with the radio frequencies used by network nodes [18]. In the event that an attacker uses a rather powerful jamming source, disruptions of networks' proper function are likely to occur.

In jamming attacks, a jammer can simply disregard the medium access protocol (MAC) by continually transmitting signal on a wireless channel. By doing so, the jammer either prevents users from being able to commerce with legitimate MAC operations, or introduces packet collisions that force repeated backoffs [19]. The objective of the jammer is to interfere with legitimate wireless communications. The jammer can achieve this goal by either preventing a real traffic source from sending out a packet, or by preventing the reception of legitimate packets. When the jamming occurs, the traffic going through the jammed area is disrupted, and the traffic needs to be rerouted around the jammed area.

We consider two rerouting strategies; global rerouting and local rerouting. They have tradeoffs between the rerouting latency and network performance after rerouting. In global rerouting, all traffic in the network will be rerouted. Local rerouting uses a set of detour paths to route around the jammed area locally. The local rerouting strategy can typically restore service much faster than the global rerouting because the restoration is locally activated. In this paper, we investigate the local rerouting strategies that can minimize the performance degradation in the event of jamming attacks.

Several approaches are proposed in recent works to address the jamming issue. Xu et al. [19] consider how to detect jamming where congested. They introduce the notion of consistency checking, where the packet delivery ratio is used to identify a radio link that has poor utility. Then signal strength consistency check is performed to classify whether the poor link quality is due to jamming. JAM (Jammed-Area Mapping) [20] focuses on the method to be used after jamming detection. It uses a priority message to inform neighbors of the attack detection, and it maps the jammed area as feedback for routing. However, it takes time for the routing protocol to update the information. During the time, normal traffic routed to the jamming area may become congested or dropped. Besides, a single detour path to a

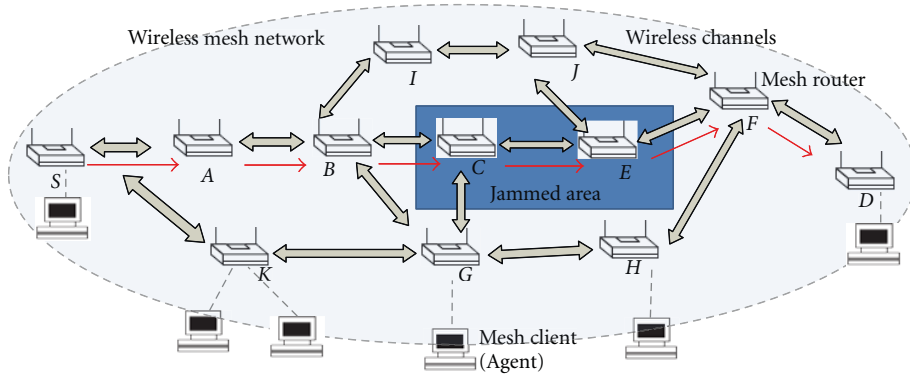


FIGURE 2: WMN system model.

destination generated by general routing protocols could easily become congested again. In this paper, we assume that the detection and mapping of jammed area are serviced by the conventional schemes [17, 20, 21]. In other words, the information of a jammed area, such as node ID in the area and locations of nodes, is fed back to nodes in the network by the service.

2.2. Architecture of WMN System. The WMN is a group of mesh clients and routers interconnected via wireless links. Mesh clients (MCs) can be various devices with wireless network interface cards such as PCs and laptops. In this paper, an agent takes a role of a MC, and hereafter we call an MC as an agent. Agents have limited communication resources and capability. Mesh routers (MRs) are usually powerful in terms of computation and communication capabilities. They normally stay static and act as access points to supply network connections for the agents. Due to limited radio coverage range and dynamic wireless channel capacity, message from an agent is usually transmitted through a multihop path to its destination. Ad hoc mode interconnections of the MRs construct the wireless mesh backbone network.

Our WMN structure is illustrated in Figure 2. On each MR, one wireless channel is assigned for communication with agents, while the other channel is assigned for communication with other MRs. When a source agent intends to send its data to a destination agent, the source agent sends the data to its neighboring MR, and the MR starts to construct a routing path towards the MR neighboring to the destination agent. Once the data arrives at the MR, neighboring to the destination and the MR forwards the data to the destination agent.

We model a multiradio multichannel WMN as a directed graph $G = (V, E, C)$, where $v \in V$ represents a WMN node (i.e., a MR) and $e \in E$ represents a wireless link between two nodes. We assume that our WMN uses a set of orthogonal wireless channels denoted by C. We consider the WMN under jamming attacks. The wireless communications of jammed nodes are interfered by RF signals or packets sent from jammers. A jammed node at channel $c \in C$ is denoted as $j_c \in J_c$, where J_c is a set of jammed nodes detected at channel c and J is the set of jammed nodes over all channels. An unjammed node is denoted as $u_c \in U_c$, where U_c is a set of unjammed nodes at channel c and U is a set of unjammed nodes over all channels where $U = V - J$. In Figure 2, nodes C and E are jammed nodes. When there is a traffic flow from source S to destination D, node B identifies that the next nodes C and E over its path are jammed by conventional jammed-area mapping services. Node B starts to discover local detour subpath to bypass the jammed area. For traffic flow f, $\text{pre}_f(J_c)$ represents a set of nodes sending data directly to one or more nodes in J_c. In Figure 2, node B is a member of $\text{pre}_f(J_c)$. The objective of our traffic rerouting scheme is to find detour path $b_f(v_{\text{pre}}, J_c)$ of flow f in the WMN, and $b_f(v_{\text{pre}}, J_c)$ is a detour path to solve the jamming problem caused by jammed nodes J_c with $v_{\text{pre}} \in \text{pre}_f(J_c)$.

3. The Proposed Traffic Rerouting Scheme

When a node $v_{\text{pre}} \in \text{pre}_f(J_c)$ detects jammed area, it constructs the detour subpath. However, if nodes to detect jamming attacks simultaneously switch to the new link, the amount of input traffic to the new link increases abruptly and congests the link. Packet loss is also increased. To solve the problem, we employ stochastic rerouting scheme to reduce the congestion and offer load balancing. In other words, a detour path is stochastically selected among multiple candidates of the detour path. To determine the candidates, we have considered multipath routing schemes in WSNs [22–28]. Our traffic rerouting scheme is composed of two steps; determining candidates of a detour subpath and selecting the detour subpath among the candidates. In this paper, we assume that all nodes in the WMN know their own location, a destination's location, and location information of the WMN system by using GPS or any localization schemes [29, 30]. In our proposed scheme, a packet sent from a source is forwarded to a destination using a conventional geographic routing scheme [31].

In the first step, we employ anchor points to construct the disjoint multiple candidates. An anchor point is a reference point to determine the next hop node over a candidate detour path. In other words, a node determines the closest node to the anchor node among its neighboring nodes as the next-hop node by using a conventional geographic routing scheme. Anchor points are divided into two groups; entry anchor points and exit anchor points. In this paper, we use k

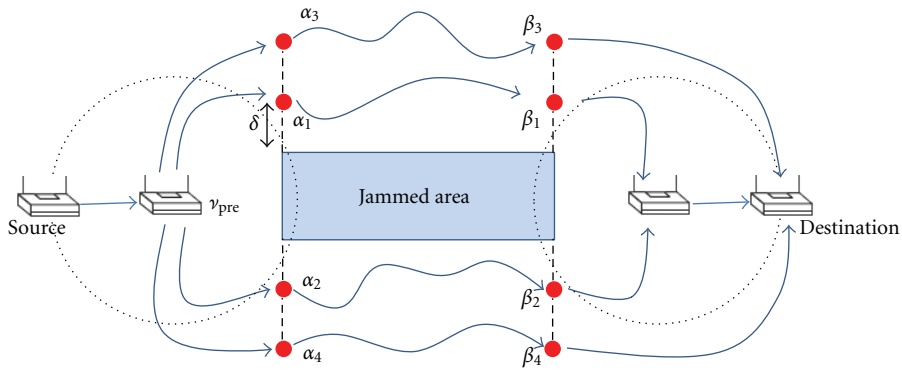


Figure 3: k multiple candidates of a detour subpath.

entry anchor points and k corresponding exit anchor points to construct k multiple candidates of the detour path. The number of candidates paths, k, is dependent on the degree of service desired by a microgrid. The separation distance between neighboring two entry points or neighboring two exit points is δ to make physically disjoint multiple candidates. As mentioned previously, jamming is the type of attack which interferes with the radio frequencies used by network nodes. Thus, in order to reduce the interference by the attack, we set δ to $R \leq \delta \leq 2R$, where R is a transmission range of a node. When v_{pre} identifies that the next node over its path is jammed, it determines entry anchor points and exit anchor points. It designates the first entry anchor point, α_1, with δ distance from the first jammed node in the jammed area. Then other entry anchor points, α_i $(1 < i \leq k)$, are designated with δ distance from the previous entry anchor point, α_{i-1}. Similarly, v_{pre} designates exit anchor points, β_1, with δ distance from the last jammed node in the jammed area. Other exit anchor points, β_i $(1 < i \leq k)$, are designated with δ distance from the previous exit anchor point, β_{i-1}. Once anchor points are determined, a 3-tuple $< v_{pre}, \alpha_i, \beta_i >$ is a candidate of the detour path. Figure 3 shows k disjoint multiple candidates.

Once k candidates of the detour subpath are determined, the second step of our traffic rerouting scheme is initiated to select the detour subpath among the candidates. When the candidates are determined, v_{pre} stochastically selects the detour path. The next nodes of v_{pre} over multiple candidates are ordered by link qualities of the next nodes, such as RSS (received signal strength), LQI (link quality indication), or SNR (signal-to-noise ratio). The link qualities of the next nodes are transformed into a value between 0 and 1 using the min-max normalization. Then, v_{pre} selects its next node with probability p, and the corresponding candidate is selected as the detour subpath.

When the detour path is selected, v_{pre} sends the data received from the source to the entry anchor point in the selected 3-tuple by using a conventional geographic routing scheme. When an intermediate node is in R range of the entry anchor point the node declares itself as an entry anchor node for the entry anchor point, and it sends the message to the corresponding exit anchor point. An exit anchor node is selected in similar way. In other words, when an intermediate node is in R range of the exit anchor point, the node declares itself as an exit anchor node for the exit anchor point, and it sends the data to the destination node. The traffic rerouting is completed.

4. Performance Evaluations and Discussions

To quantitatively evaluate the performance of the proposed scheme, we use NS-2 network simulator [32]. Our WMN delivers the microgrid-related data between agents. In a 1500 m by 1500 m grid, we deploy 30 MCs, that is, agents and 70 MRs. The MCs are randomly placed and MRs are equally spaced. MCs exchange 1024-byte CBR packets with other randomly selected MCs through MRs every second. As comparative routing mechanism, JAM [20] with a detour path is simulated. As performance metrics, PDR (Packet Delivery Ratio) and end-to-end delay of normal traffic are measured. PDR is the ratio of total number of packets received by the nodes to the total number of packets transmitted. End-to-end delay is the time taken for a packet to be transmitted across a network from source to destination. For the simulation parameters, we vary the number of normal flows and the number of nodes within the jammed area. The jammed area is located at the center of network; thus generated normal traffic would be destined for the jammed area in normal routing. In our experiments, the transmission range of a node is defined as 250 m, and two-ray ground propagation channel is assumed with a data rate of 54 Mbps. We used IEEE 802.11 g MAC, and the total simulation time is 3600 seconds. In figures, k indicates the number of candidates subpath in our scheme ("*PS*") and "*normal*" indicates that the network has no jamming traffic.

Figures 4 and 5 show PDR and end-to-end delay when the number of normal flows increases. The number of nodes in the jammed area is set to 4. JAM provides a high PDR until the number of normal flows is 5, but PDR decreases in the case with a lot of normal traffic, and this induces a high end-to-end delay. Since JAM selects the best single path to detour the jammed area, the selected path easily becomes congested when the number of normal flows increases. In comparison with JAM, our scheme increases PDR by about 10% and decreases the end-to-end delay by about 80% because of the stochastical traffic distribution among

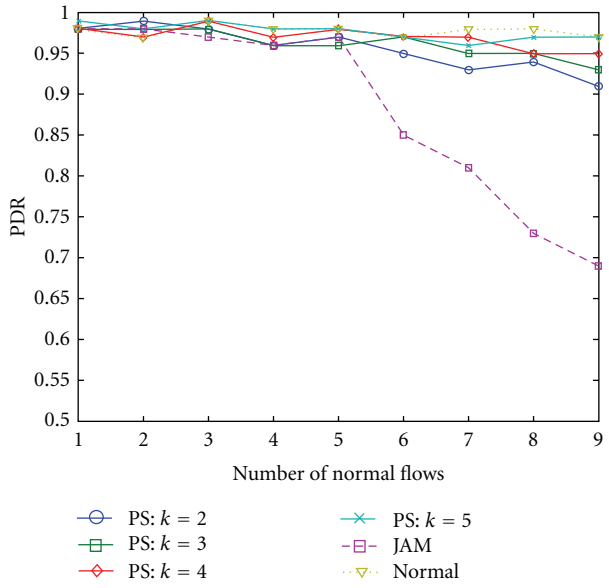


Figure 4: Performance comparison of PDR with varying the number of normal flows.

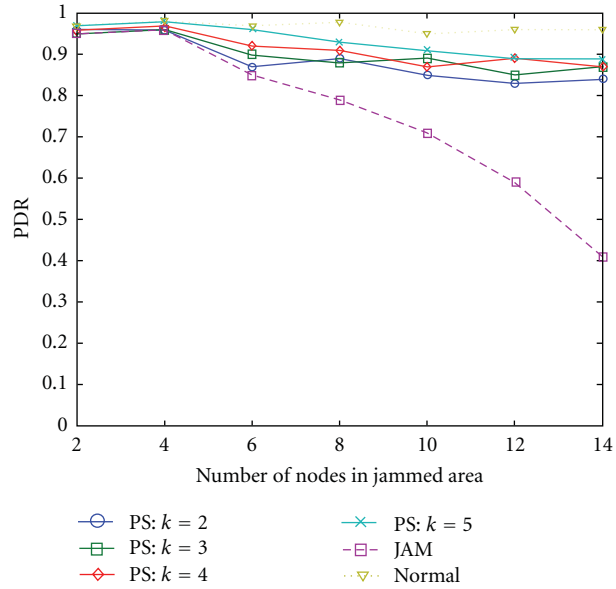


Figure 6: Performance comparison of PDR with varying the number of nodes in jammed area.

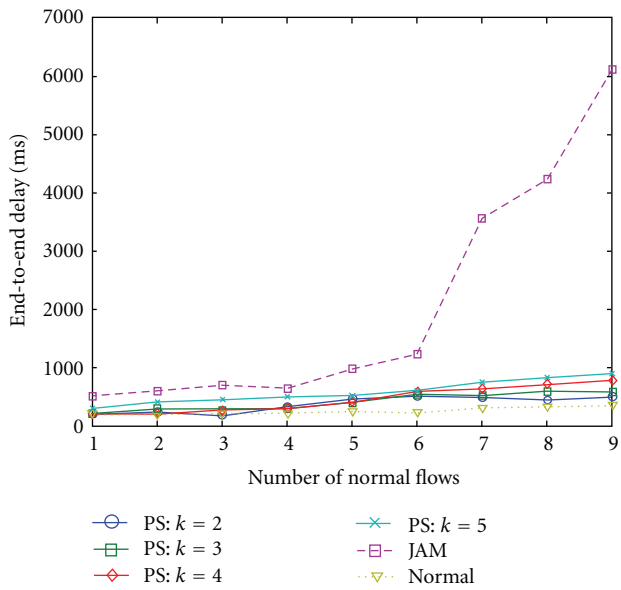


Figure 5: Performance comparison of the end-to-end delay with varying the number of normal flows.

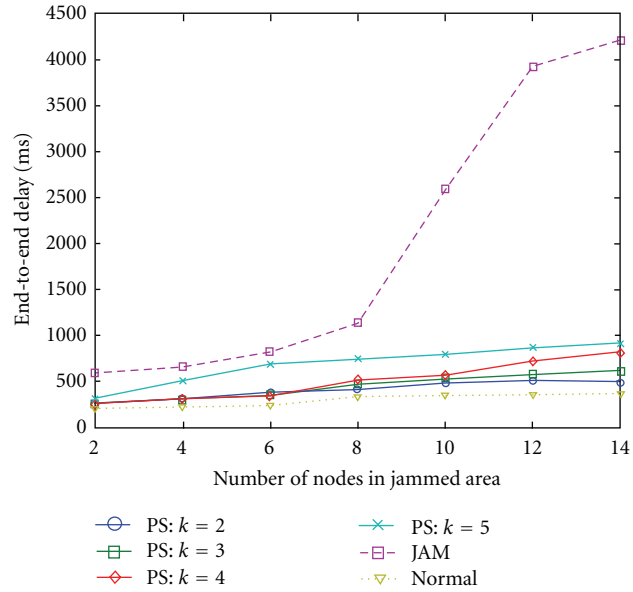


Figure 7: Performance comparison of the end-to-end delay with varying the number of nodes in jammed area.

multiple candidates. In our scheme, when k increases, the distance between the candidate path and the jammed area increases and the effect of the jamming attack decreases. Thus, PDR increases. However, when the distance increases, the length of the candidate subpath increases and the delay also increases.

Figures 6 and 7 show PDR and end-to-end delay when the number of nodes in the jammed area increases. The number of normal flows is set to 4. PDR and the end-to-end delay of JAM are worsened drastically in the big jammed area because of traffic concentration on a long detour path. However, our scheme stochastically distributes normal flows on disjoint multiple paths and efficiently balances the traffic load. Thus, we can see that our scheme achieves an improvement of about 20% and 75% in PDR and the end-to-end delay, respectively. For our scheme, we carried

out additional experiments to address a scalability problem. From the results, we can conclude that our scheme is scalable when the number of nodes increases from 2 to 50. For example, PDR and end-to-end delay are 78% and 1.42 sec, respectively, when the number of nodes is 50.

5. Conclusion

In this paper, we focused a traffic rerouting scheme under jamming attacks in wireless communication infrastructure for islanded microgrid. In conventional schemes, when nodes detect jamming attacks, they simultaneously switch to a new link, and the amount of input traffic to the new link increases abruptly and congests the link. To solve the problem, we employ a stochastic rerouting scheme to reduce the congestion and offer load balancing. First, we determine physically disjoint multiple paths as the candidates of a detour path to reduce the effect of jamming attack. Then, we stochastically select the detour path among the multiple candidates to offer load balancing. From the performance comparison, we show that the performance of our scheme is better than that of conventional scheme in terms of the packet delivery ratio and the end-to-end delay.

In our scheme, one of the factors affecting the performance is the distance δ. To optimize the performance, we need a dynamic decision scheme of δ in time-varying environments. We will consider developing the decision scheme as a future research direction.

Acknowledgment

This work was supported by the GRRC program of Gyeonggi province ((GRRC SUWON2011-B4), Research on Precision Location Tracking System for Real-Time Situation).

References

[1] B. Lasseter, "Role of distributed generation in reinforcing the critical electric power infrastructure," in *IEEE Power Engineering Society Winter Meeting*, pp. 146–149, February 2001.

[2] J. Y. Kim, S. K. Kim, and J. H. Park, "Contribution of an energy storage system for stabilizing a microgrid during islanded operation," *Journal of Electrical Engineering and Technology*, vol. 4, no. 2, pp. 194–200, 2009.

[3] J. Y. Kim, J. H. Jeon, S. K. Kim et al., "Cooperative control strategy of energy storage system and microsources for stabilizing the microgrid during islanded operation," *IEEE Transactions on Power Electronics*, vol. 25, no. 12, pp. 3037–3048, 2010.

[4] J. H. Jeon, J. Y. Kim, H. M. Kim et al., "Development of hardware in-the-loop simulation system for testing operation and control functions of microgrid," *IEEE Transactions on Power Electronics*, vol. 25, no. 12, pp. 2919–2929, 2010.

[5] H. M. Kim and T. Kinoshita, "A new challenge of microgrid operation," *Communications in Computer and Information Science*, vol. 78, pp. 250–260, 2010.

[6] A. L. Dimeas and N. D. Hatziargyriou, "Operation of a multiagent system for microgrid control," *IEEE Transactions on Power Systems*, vol. 20, no. 3, pp. 1447–1455, 2005.

[7] A. L. Dimeas and N. D. Hatziargyriou, "Agent based control for microgrids," in *IEEE Power Energy Society General Meeting*, June 2007.

[8] H. M. Kim, T. Kinoshita, Y. Lim, and T. H. Kim, "A bankruptcy problem approach to load-shedding in multiagent-based microgrid operation," *Sensors*, vol. 10, no. 10, pp. 8888–8898, 2010.

[9] H. M. Kimt and T. Kinoshita, "A multiagent system for microgrid operation in the grid-interconnected mode," *Journal of Electrical Engineering and Technology*, vol. 5, no. 2, pp. 246–254, 2010.

[10] H. M. Kim, T. Kinoshita, and M. C. Shin, "A multiagent system for autonomous operation of islanded microgrids based on a power market environment," *Energies*, vol. 3, no. 12, pp. 1972–1990, 2010.

[11] H. M. Kim, T. Kinoshita, and Y. Lim, "Talmudic approach to load shedding of islanded microgrid operation based on multiagent system," *Journal of Electrical Engineering and Technology*, vol. 6, no. 2, pp. 284–292, 2011.

[12] H.-M. Kim, W. Wei, and T. Kinoshita, "A new modified CNP for autonomous microgrid operation based on multiagent system," *KIEE Journal of Electrical Engineering & Technology*, vol. 6, pp. 139–146, 2011.

[13] M. Wooldridge, *An Introduction to Multiagent Systems*, John Wiley and Sons, New York, NY, USA, 2nd edition, 2009.

[14] OpenSG, "SG Network System Requirements Specification r3," May 2010.

[15] F. Huang, Y. Yang, and L. He, "A flow-based network monitoring framework for wireless mesh networks," *IEEE Wireless Communications*, vol. 14, no. 5, pp. 48–55, 2007.

[16] S. Jiang and Y. Xue, "Providing survivability against jamming attack for multi-radio multi-channel wireless mesh networks," *Journal of Network and Computer Applications*, vol. 34, no. 2, pp. 443–454, 2011.

[17] A. Mpitziopoulos, D. Gavalas, C. Konstantopoulos, and G. Pantziou, "A survey on jamming attacks and countermeasures in wsns," *IEEE Communications Surveys and Tutorials*, vol. 11, no. 4, pp. 42–56, 2009.

[18] E. Shi and A. Perrig, "Designing secure sensor networks," *IEEE Wireless Communications*, vol. 11, no. 6, pp. 38–43, 2004.

[19] W. Xu, W. Trappe, Y. Zhang, and T. Wood, "The feasibility of launching and detecting jamming attacks in wireless networks," in *6th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MOBIHOC '05)*, pp. 46–57, May 2005.

[20] A. D. Wood, J. A. Stankovic, and S. H. Son, "JAM: a jammed-area mapping service for sensor networks," in *24th IEEE International Real-Time Systems Symposium (RTSS '03)*, pp. 286–297, December 2003.

[21] R. Muraleedharan and L. Osadciw, "Jamming attack detection and countermeasures in wireless sensor network using ant system," in *SPIE Symposium on Defense and Security*, Orlando, Fla, USA, April 2006.

[22] H. W. Oh, J. H. Jang, K. D. Moon, S. Park, E. Lee, and S. H. Kim, "An explicit disjoint multipath algorithm for cost efficiency in wireless sensor networks," in *21st International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC '10)*, pp. 1899–1904, September 2010.

[23] D. Ganesan, R. Govindan, S. Shenker, and D. Estrin, "Highly-resilient, energy-efficient multipath routing in wireless sensor networks," *ACM Mobile Computing and Communications Review*, vol. 1, pp. 10–24, 2002.

[24] E. P. C. Jones, M. Karsten, and P. A. S. Ward, "Multipath load balancing in multi-hop wireless networks," in *IEEE*

International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob '05), pp. 158–166, August 2005.

[25] S. Waharte and R. Boutaba, "Totally disjoint multipath routing in multihop wireless networks," in *IEEE International Conference on Communications (ICC '06)*, pp. 5576–5581, July 2006.

[26] E. Gelenbe and E. C. H. Ngai, "Adaptive qos routing for significant events in wireless sensor networks," in *5th IEEE International Conference on Mobile Ad-Hoc and Sensor Systems (MASS '08)*, pp. 410–415, October 2008.

[27] E. Gelenbe and E. C.-H. Ngai, "Adaptive random re-routing for differentiated QoS in sensor networks," in *Proceedings of the Visions of Computer Science Conference*, pp. 343–354, September 2008.

[28] E. Gelenbe and E. C.-H. Ngai, "Adaptive random re-routing in sensor networks," in *Annual Conference of ITA (ACITA '08)*, pp. 348–349, London, UK, September 2008.

[29] A. Islam, U. Iqbal, J. M. P. Langlois, and A. Noureldin, "Implementation methodology of embedded land vehicle positioning using an integrated gps and multi sensor system," *Integrated Computer-aided Engineering*, vol. 17, no. 1, pp. 69–83, 2010.

[30] F. G. Bravo, A. Vale, and M. I. Ribeiro, "Navigation strategies for cooperative localization based on a particle-filter approach," *Integrated Computer-aided Engineering*, vol. 14, no. 3, pp. 263–279, 2007.

[31] B. Karp and H. T. Kung, "Gpsr: greedy perimeter stateless routing for wireless networks," in *6th Annual International Conference on Mobile Computing and Networking (MOBICOM '00)*, pp. 243–254, August 2000.

[32] "The network Simulator ns-2," http://www.isi.edu/nsnam/ns/.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 954398, 8 pages
doi:10.1155/2012/954398

Research Article

A Network Coding Based Rarest-First Packet Recovery Algorithm for Transmitting Geocast Packets over Hybrid Sensor-Vehicular Networks

Tz-Heng Hsu and Ying-Chen Lo

Department of Computer Science and Information Engineering, Southern Taiwan University, Tainan, Taiwan

Correspondence should be addressed to Tz-Heng Hsu, hsuth@mail.stut.edu.tw

Received 27 October 2011; Revised 8 January 2012; Accepted 11 January 2012

Academic Editor: Tai Hoon Kim

Vehicular ad hoc network (VANET) and road-side sensors are used to improve driving safety in many applications. Sensor nodes deployed along the roadside are used to sense road conditions and then deliver information about dangerous conditions to vehicles. In hybrid sensor-vehicular networks, new challenges arise and should be addressed. Geocasting can be used to perform the regional broadcast to deliver geographic-related safety, commercials, and advertisements messages. The challenging problem in geocasting is how to deliver packets to all the nodes within the geocast region with high efficiency but low overhead. Network coding is a special in-network data-processing technique that can potentially increase the network capacity and packet throughput in wireless networking environments. In this paper, a network coding based rarest-first packet recovery algorithm for transmitting geocast packets over hybrid sensor-vehicular-networks is proposed. The proposed algorithm can increase packet delivered ratio at each mobile node. As a result, the safety and transmission efficiency can be achieved simultaneously.

1. Introduction

A vehicular ad hoc network (VANET) provides communications among nearby vehicles and between vehicles and fixed equipment, for example, road-side unit (RSU). People can access telematic services through the road-side unit in VANETs. With numerous connectivity facilities for vehicles, telematic services have become extremely popular, and its use grows exponentially. To provide vehicle safety services, such as emergency warning system, lane-changing assistant, and anticollision system, roadside-sensor nodes are often deployed along the road to sense and collect road conditions. A hybrid sensor-vehicular network combines wireless communications provided by vehicular ad hoc networks (VANETs) with sensing devices in cars and road-side sensor infrastructure, that is, wireless sensor networks (WSNs). The roadside-sensor nodes can form a wireless sensor network to communicate with each other through the wireless interface. When a vehicle accident occurs, the sensors can deliver warning messages to vehicles for safe driving.

Geocasting is a kind of regional broadcasting, which sends messages to a number of nodes within a specific region. Geocasting can be used to perform the regional broadcast to deliver geographic-related safety, commercial, and advertisement messages. For telematic applications in VANETs, geocasting is an important and frequent communication service. Geocasting is widely used to deliver packets to nodes within a certain geographical region for sending emergency messages. In a wireless sensor network, nodes may take on different roles, that is, data source nodes, intermediate nodes, and sink nodes, according to their tasks. The roadside-sensor nodes can use the geocasting scheme to deliver the sensed data to vehicles within a certain geographical region to provide the drivers with early warnings and avoid the accident from occurring again. The challenging problem in geocasting is how to deliver the packets to all the nodes within the geocast region with high efficiency but with low overhead [1–3].

A source node may reside inside or outside the geocast region. When the source node resides inside the geocast

region, the source node floods the packets to all nodes within the geocast region for disseminating location-based information. When the source node resides outside the geocast region, a forwarding zone (FZ) is defined as a partitioned network between the source node and the geocast region. The size of the forwarding zone is determined by (i) the size of the geocast region and (ii) the location of the source node. The current location is determined using GPS receivers at each node. When a node outside the geocast region receives a geocast packet and the packet was not received previously, the node will forward the packet to its neighbors if it is belonged to the forwarding zone; otherwise, it will discard the packet. Once a geocast packet is reached at the geocast region and a node is received, the node will accept the packet. The node will also broadcast the packet to its neighbors, if it has not received the packet previously (repeated reception of a packet is checked using sequence numbers in the packet header).

Network coding is a special in-network data-processing technique that can potentially increase the capacity and the throughput of the wireless network. Network coding has recently emerged as a promising generalization by allowing nodes to mix and then forward information received on its incoming links. With network coding, intermediate nodes may send out packets that are linear combinations of previously received information. The main benefits of network coding are (1) network throughput improvements and (2) robustness enhancement on data transmission.

In a sensor network, geocasting may be required to query node conditions in a certain area. In VANET, geocasting can be used to send emergency warnings to a region. In urban areas, VANET may suffer from greatly high node density. To increase the throughput of geocast packets that reached at the geocast region with high node density, the multiforwarding zone concept, which is our previous work depicted in [4], can be adopted to increase the throughput of geocast packets at the geocast region with high node density. In this paper, a network coding based rarest-first packet recovery algorithm for transmitting geocast packets over hybrid sensor-vehicular networks is proposed to increase packet delivered quality at each vehicle. By equipping network coding technique for geocast packets according to network-status, the proposed algorithm can enhance the utilization of packet transmission and can satisfy the requirement of adaptive telematic services over hybrid sensor-vehicular networks.

The rest of this paper is organized as follows: Section 2 introduces related works about geocasting and network coding. Section 3 describes the proposed network coding based transmission architecture for delivering geocast packets over hybrid sensor-vehicular networks. Section 4 shows the experimental results. Section 5 has the conclusion remarks.

2. Related Works

In this section, works related to geocasting and network coding are presented.

2.1. Geocasting. Geographical routing requires no routing information exchange except the discovery of the positions of the neighbors within a single hop. Geographical routing has been applied to geocast and becomes a promising routing method for wireless sensor networks due to its simplicity. Geocast protocols can be mainly categorized based on flooding the network or on forwarding a geocast packet on a particular routing path [1]. In [2], Ko and Vaidya presented two different location-based multicast schemes to decrease delivery overhead of geocasting packets. The proposed algorithms limit the forwarding space for transmitting packets to the forwarding zone. Simulation results show that the proposed schemes can reduce the message overhead significantly. Meanwhile, it is possible to achieve accuracy of multicast delivery comparable with multicast flooding. In [3], two location-aided routing (LAR) protocols are proposed to decrease the overhead of route discovery by utilizing location information for mobile hosts. The LAR protocols use location information to reduce the search space for a desired route, which results in fewer route discovery messages. In [4], Hsu et al. proposed geocasting algorithms using Voronoi diagram to forward messages to neighbors who may be the best choices for a possible position of destination. The V-GEDIR routing method determines neighbors that may be closest to the destination. The CH-MFR is based on the convex hull on neighboring nodes. The proposed V-GEDIR and CH-MFR algorithms are loop-free and have smaller flooding rate compared to other directional geocasting methods.

In [5], Stojmenovic et al. proposed three mesh approaches for delivering packets to the geocast group: FLOOD, BOX, and CONE. The mesh provides redundant paths between the source and the destination region against host mobility and link failures. In the proposed mesh approaches, the mesh is created in initial step for discovering redundant paths. Nodes inside the destination region will join the mesh while receiving the initial packet. After an unicast reply is sent back to the sender according to the reverse path, the flooding is stopped. The FLOOD approach has the highest control overhead and network-wide data load, but provides the highest level of reliability. The CONE approach has the smallest control overhead and network-wide data load, but provides the smallest level of reliability. The BOX approach fall between the FLOOD and CONE approaches. In a BOX forwarding zone, the forwarding zone is defined as the smallest rectangle that covers both the source node and the geocast region. In a CONE forwarding zone, the forwarding zone is defined as a cone rooted at source node, such that angle made by the cone is large enough to include the forwarding zone [2]. Consider a source node S that needs to send geocast packets to all nodes that are currently located within a certain geographical region, that is, geocast region. Reducing the area of the forwarding zone can reduce the control overhead and network-wide data load. However, a geocast packet may be lost due to no node exists in the reduced area of the forwarding zone. In a CONE forwarding zone, such a case occurs when the angle made by the cone is not large enough.

A geocast packet will be lost if no enough nodes in the forwarding zone (FZ) can forward the packet to the geocast region. Figure 1 shows an example of packet loss caused by no enough nodes in the forwarding zone (FZ). In Figure 1, node A, B, and C cannot receive the geocast packets from source

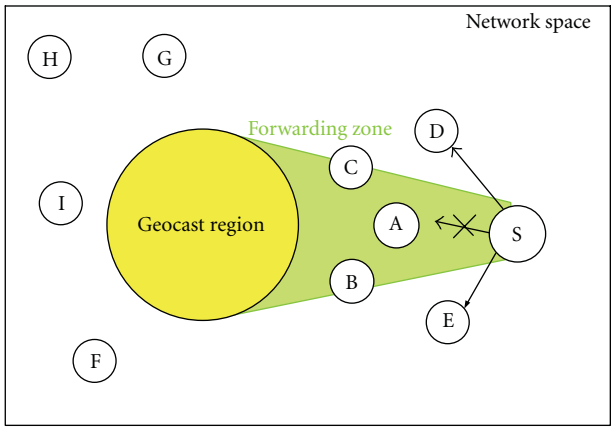


Figure 1: An example of packet loss caused by no enough nodes in the forwarding zone (FZ).

node S due to that no enough nodes can forward packet to them. Node A cannot receive packets from source node S due to channel losses induced by the long-distance transmission. Although node D and E can receive packets from source node S, however, node D and E drop the received packets and will not forward the packets to other nodes, for example, B and C, because of node D and E are not belonging to the forwarding zone. Forwarding zone adaptation can solve the above-mentioned issues. In a CONE forwarding zone, the forwarding zone can be expanded into a BOX forwarding zone to avoid packet loss caused by the small angle made by the cone. In a BOX forwarding zone, a parameter δ can be used to extend the forwarding zone [2]. When δ is positive, the rectangular forwarding zone is extended. In case of still no node exists in the extended forwarding zone, the source node can use the simplest flooding mechanism to flood geocast packets to the destined geographical region. Source node broadcasts the geocast packet to all its neighbors. The flooding mechanism can increase the possibility that packets can be relayed to the destined geocast region. Meanwhile, the flooding mechanism is simple and easy to implement. However, it will increase the transmission overhead and network-wide data load as a tradeoff between transmission efficiency and packet reachability.

In [6, 7], Ko and Vaidya presented a novel protocol called GeoTORA for geocasting packets in ad hoc networks. The GeoTORA protocol is obtained using a small variation on the TORA anycasting protocol. In GeoTORA, flooding is limited to nodes within a small region. GeoTORA can significantly reduce the overhead of geocast delivery. In [8], Ko and Vaidya proposed two novel geocasting protocols for achieving high delivery rate and low overhead by utilizing the local location information of nodes with region flooding: Geographic-Forwarding-Geocast (GFG) and Geographic-Forwarding-Perimeter-Geocast (GFPG). Simulation results show that the CFG and CFPG have significant improvement in delivery rate and reduction in overhead. However, the authors do not consider the packet loss caused by broadcast storm within the geocast region.

Our transmission scheme differs from above-mentioned works is that we try to (1) increase the packet delivery ratio with multiforwarding zone and (2) recover the packet loss caused by broadcast storm within the geocast region at high node density using network coding.

2.2. Network Coding. Network coding, which was introduced by Seada and Helmy [9], allows each node in a network to perform some computations, for example, encoding and decoding, to enhance bandwidth utilization in network communications. Ahlswede et al. introduced network information flow problem that can be regarded as the max-flow min-cut theorem for network transmission. By employing coding at the nodes, that is, network coding, bandwidth can in general be saved. A node can function as an encoder in the sense that it receives messages from all the input links; the node encodes, and then forwards messages to all the output links. In [10], Ahlswede et al. presented an explicit construction of a code for multicast in a network that achieves the max-flow bound on the information transmission rate. Linear coding regards a block of data as a vector over a certain base field and allows a node to apply a linear transformation to a vector before passing it on. The authors formulate a multicast problem and prove that linear coding suffices to achieve the optimum, which is the max-flow problem from the source to each receiving node. The code is linear, which makes encoding and decoding easy to implement in practice. In [11], Li et al. proposed decentralized algorithms that compute minimum-cost subgraphs for establishing multicast connections in networks using network coding. Two decentralized algorithms are described for computing minimum-cost subgraphs: one that applies for linear cost function and the other that applies for strictly convex cost functions. The proposed decentralized approach can achieve minimum-cost multicast with network coding.

In [12], Lun et al. introduced an opportunistic approach to network coding named COPE, where each node snoops on the medium, obtains the status of its neighbors, detects coding opportunities, and codes as long as the recipients can decode. COPE can efficiently support multiple unicast flows even when traffic demands are unknown and bursty. Emulation results show that COPE substantially improves the network throughput. COPE's throughput becomes many times higher than current 802.11 mesh networks as the number of flows and the contention level increases. In [13], Katti et al. proposed a new protocol named CAPO for multihop wireless networks. In addition to code packets, CAPO attaches nonencoded packets together opportunistically. CAPO has a lazy decoding process which tries to decode stored coded packets to increase the decoding probability, which can reduce the number of retransmissions using the minimum number of decodings. In comparison with COPE, CAPO with its opportunistic attaching and lazy decoding feature can reduce the number of transmissions noticeably.

3. The Network Coding Based Rarest-First Packet Recovery Algorithm

In urban areas, VANET may suffer from greatly high node density. To increase the throughput of geocast packets that reached at the geocast region with high node density, we

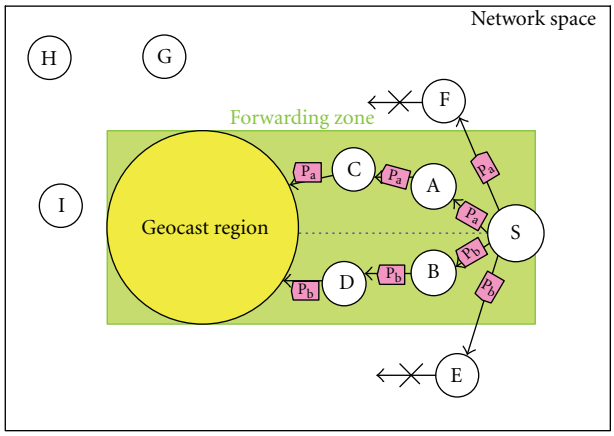


Figure 2: An example of delivering geocast packets using a BOX multiforwarding zone (MFZ).

propose a geocasting mechanism using the multi-forwarding zone concept [4]. The proposed multi-forwarding zone (MFZ) geocasting mechanism splits the original forwarding zone into two or more small forwarding zones. The simplest way is to draw a virtual cutting line from the center of the geocast region to the source node. The virtual cutting line splits the original forwarding zone into two small forwarding zones: upper forwarding zone (UFZ) and lower forwarding zone (LFZ). The size of the upper forwarding zone (UFZ) and lower forwarding zone (LFZ) are the same. For nodes that are equipped antennas with multiinput multioutput (MIMO), geocast packets can be forwarded simultaneously using the upper forwarding zone (UFZ) and lower forwarding zone (LFZ). For nodes with single antenna, geocast packets can be forwarded interleavedly using the upper forwarding zone (UFZ) and lower forwarding zone (LFZ).

In case of that no enough nodes exist in the extended multiforwarding zone (MFZ), forwarding zone adaptation policy is adopted to increase the possibility that packets can be relayed to the destined geocast region. In a CONE multiforwarding zone, the multiforwarding zone can be expanded into a BOX multiforwarding zone to avoid packet loss caused by the small angle made by the cone. In a BOX multiforwarding zone, we adopt the parameter δ to extend the multiforwarding zone as listed in [2]. Figure 2 shows an example of delivering geocast packets using a BOX multiforwarding zone (MFZ). In Figure 2, the upper forwarding zone (UFZ) is responsible for delivering geocast packet Pa; the lower forwarding zone (LFZ) is responsible for geocast packet Pb. The throughput of geocast packets that reached at the geocast region will be increased by using the proposed multiforwarding zone (MFZ) geocasting mechanism. If there still no node exists in the extended multiforwarding zone (MFZ), the source node will use the simplest flooding mechanism to flood geocast packets to the destined geographical region. The flooding mechanism can increase the possibility of packets that can be relayed to the destined geocast region.

Using the multiforwarding zone (MFZ) geocasting mechanism can transmit packets to the geocast region quickly and efficiently. However, it will cause the broadcast storm when too many geocast packets are arrived at the destined geocast region in a certain period. In the geocast region, when multiple nodes receive a geocast packet, these nodes will all rebroadcast the packet to its neighbors if the packet has not been received before. When multiple nodes re-broadcast the packet to its neighbors simultaneously, it will make the severe contention on the radio channel. Blindly flooding induces the broadcast storm, which wastes precious bandwidth by sending redundant packets that will probably collide.

In order to recover the packet loss caused by broadcast storm within the geocast region, we adopted the network coding technique to reduce the number of retransmissions using the minimum number of network-coded packets. Each node generates a new packet, which is a linear combination of the earlier received packets on the link, by coefficients in the finite field. A Galois field GF(q), that is, field with a finite number of elements q, is adopted to encode packets. We will use Galois fields for $q = 2^m$, where m is 5 in our algorithm. Each packet is represented as a binary string of m bits and will perform XOR operations (multiplications and additions). All the coefficients used are randomly selected and will be embedded in the header of the network-coded packet.

In order to enhance the network coding performance, we propose a network coding based rarest-first packet recovery algorithm for transmitting geocast packets over hybrid sensor-vehicular networks. When a node receives a recovery request packet, the node puts the request packet into a temporary buffer. The node then triggers a lazy recovery timer for the first received request packet, which is used to gather more request information of its neighbor nodes. When the lazy recovery timer is timeout, for example, three seconds, the node will send a network coded recovery packet to its neighbors for recovering missing packets. The coefficients of the network coded recovery packet are determined by the network coding based rarest-first packet recovery algorithm.

The network coding based rarest-first packet recovery algorithm calculates the number of which packets are missing around its neighbor nodes. A *Recover_Rate* variable is defined for calculating the packet recovery rate of sending a specific network-coded packet. In the higher *Recover_Rate* value, the specific network-coded packet is rarest and hence can recover more missing packets for its neighbor nodes. Despite the *Recover_Rate* value, the algorithm also considers whether the network-coded packet is decodable or not at the requesting nodes. The rarest-first packet recovery algorithm employs an iterative approach known as a *level-wise* search that starts with large 1-packet sets, that is, the packet sets that contain only 1 packet and then increased the number of packet to k to discover large k-packet sets for packet recovery. Figure 3 illustrates the flowchart of the proposed network coding algorithm. The detail steps of the proposed network coding based rarest-first packet recovery algorithm are depicted as follows.

Step 1. Record the information of packets owned by source node S, and those packets are missing by S's neighbor nodes. Find the frequent missing packet number and put it into the 1-packet sets, L_1.

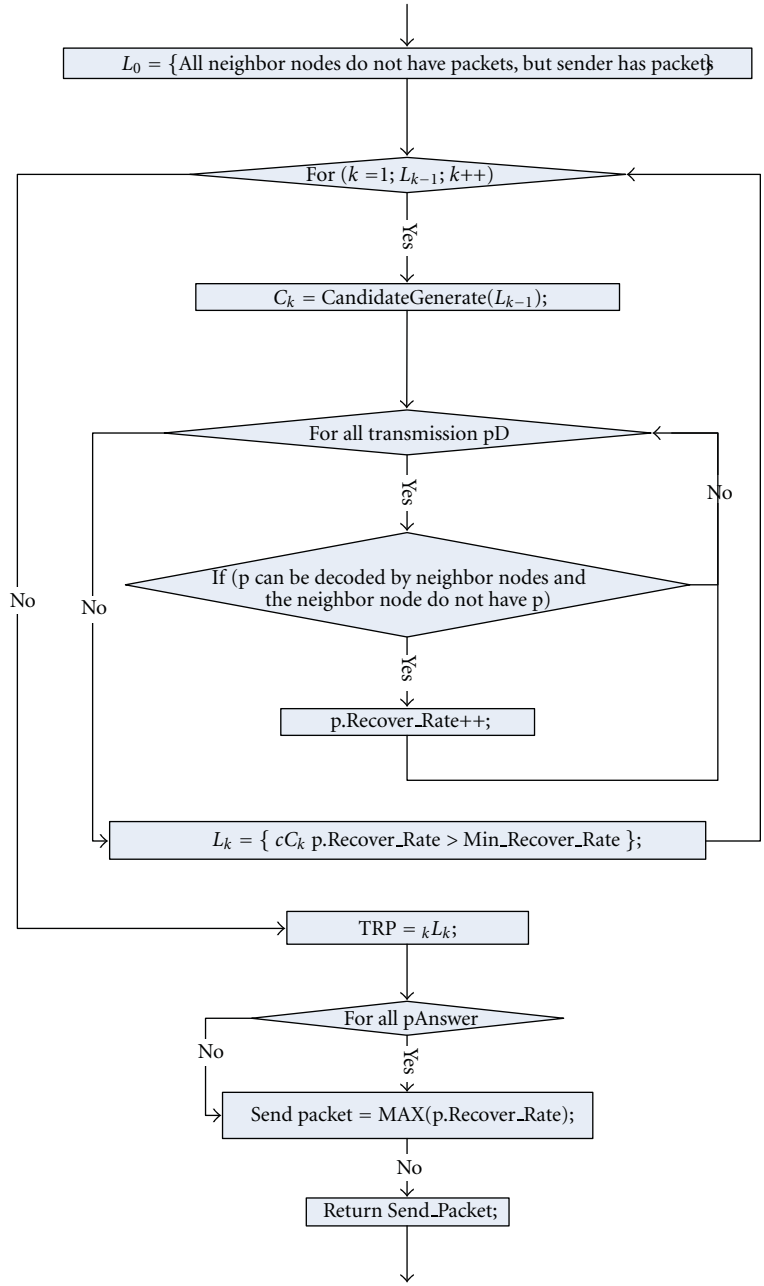


Figure 3: The flowchart of the network coding based rarest-first packet recovery algorithm.

Step 2. Calculate the *Recover_Rate* of each packet sets in packet sets L_k.

Step 3. If a packet's *Recover_Rate* is big or equal to the threshold *Min_Recover_Rate*, then the packet is put into the candidate packet sets C_k.

Step 4. If $k > 3$, then the process jumps to Step 5; otherwise, $k = k + 1$ and the process jumps to Step 2.

Step 5. Put all packets from the packet sets L_k into the packet sets transmission packet table (TRP). Find the packet p with maximum *recover_rate*.

Step 6. Once the packet p with maximum *Recover_Rate* has been found, the node encodes the packet p with network coding and then broadcasts the packet to its neighbor nodes for packet recovery.

Figure 4 shows an example of packet recovery using network coding with rarest-first packet recovery algorithm. Assume that the sender has all of the missing packets of node A, B, C, and D; node A lacks the packet 1, 3, and 4; node B lacks the packet 2, 4, and 5; node C lacks the packet 2 and 5; node D lacks the packet 2 and 5. Initially, the rarest-first packet recovery algorithm calculates the packet recovery rate for each missing packet. For example, if the sender broadcasts

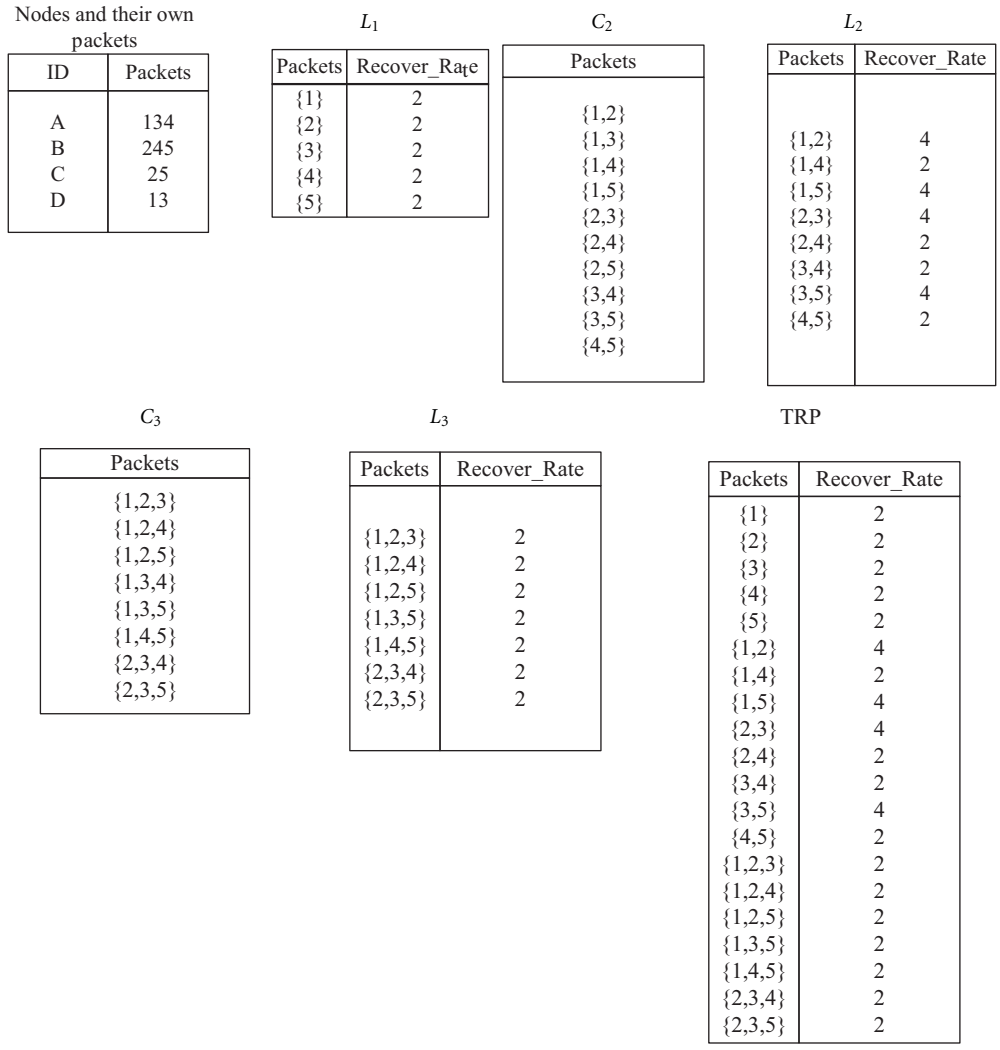


Figure 4: Packet recovery using network coding with rarest-first packet recovery algorithm.

a packet 1 to its neighbors, the value of *Recover_Rate* is calculated as 2 because node A and D can recover missed packet 1. In this example, the value of *Min_Recover_Rate* is set to 2. Since each missed packet's *Recover_Rate* is bigger or equal then to *Min_Recover_Rate,* these missed packets are put into the packet sets L_1. Then, the algorithm uses L_1 to generate candidate missed packet set C_2 in order to find L_2. The algorithm generates candidate missed packet sets and eliminates those having a subset with lower packet recovery rate, that is, *Min_Recover_Rate*. Two actions that the algorithm performs are join and prune. In the join action, L_{k-1} is joined with L_{k-1} to generate potential candidates. The prune action removes candidates that have a subset with lower packet recovery rate *Min_Recover_Rate*. Once the higher packet recovery packet sets have been found, the algorithm encodes the packet set with linear network coding and then sends the network-coded packet to its neighbors for recovering missed packets. In this example, there are four packet sets $\{1,2\}\{1,5\}\{2,3\}\{3,5\}$ with maximum packet recovery rate 4 in the TRP (transmission packet table). Since each network coded packet from these packet sets can recover 4 packets, the algorithm randomly selects a packet set for generating a network-coded packet and then broadcasts the network-coded packet to its neighbor for packets recovery.

4. Performance Evaluations

To evaluate the performance of the proposed algorithms, we use the NS2 network simulator to simulate the geocast environment. There are 19 mobile nodes in a 5000 meter by 5000 meter grid in the simulation environment; 9 mobile nodes are within the forwarding zone; 10 mobile nodes are within the geocast region. We generated a scenario file with 1 source node, which sends out a 512 bytes packet per 10 ms using constants bit rate (CBR) generator. The parameters in the simulated IEEE802.11b WLAN environment are based on Orinoco 802.11b Card [14]. The simulation runs with durations of 30 seconds. To simulate the mobility of mobile nodes, two nodes will move inside the geocasting region at

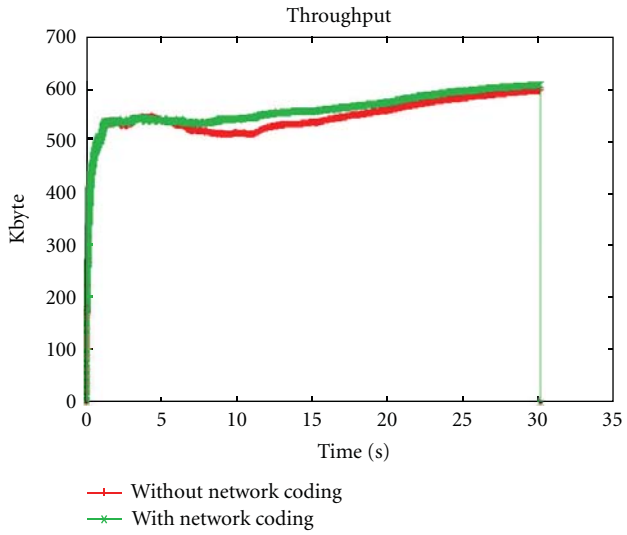


Figure 5: The results of delivering geocast packets using multiforwarding zone (MFZ) mechanisms with 10 nodes within the geocast region.

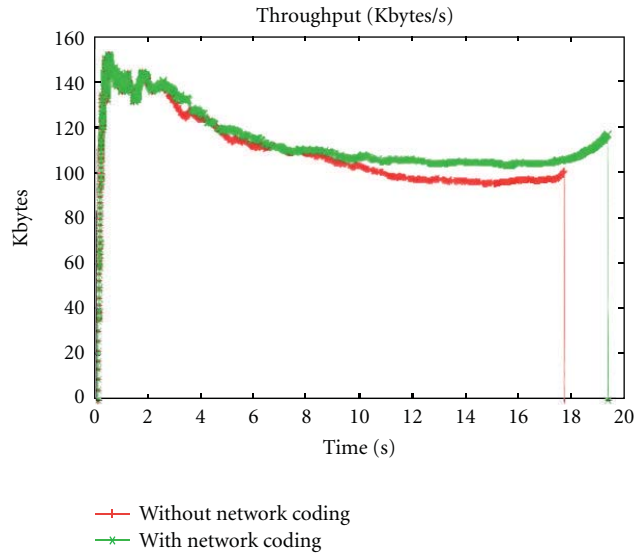


Figure 6: The results of delivering geocast packets using multiforwarding zone (MFZ) mechanisms with 30 nodes within the geocast region.

time 10 and 15. A node will move out the geocast region during time 13 to 20. Nodes within geocast region will perform network coding if the nodes receive the packet recover request from its neighbors; the packet recover request will be put into a recover request queue. A node will perform the network coding after 0.3 seconds, which can gather more recover requests. The node then codes as many request packets as the recipients can recover the lost packets. Comparisons of multiforwarding zone (MFZ) mechanisms with/without network coding are performed.

Figure 5 shows the results of delivering geocast packets using multiforwarding zone (MFZ) mechanisms with 10 nodes within geocast region. The x axis denotes the simulation time, and the y axis denotes the average packet throughput in Kbytes. The red line denotes the simulation result of delivering geocast packets using multiforwarding zone (MFZ) mechanism without network coding, the green-line denotes the simulation result of delivering geocast packets using the multiforwarding zone (MFZ) with network coding. In Figure 5, the source node sends geocast packets to the geocast region. When nodes in the geocast region receive the geocast packets, these nodes then re-broadcast the geocast packets to its neighbors. Blindly flooding induces the broadcast storm, which wastes precious bandwidth by sending redundant packets that will probably collide. The packet throughput increases rapidly at beginning and then drops at time 5 due to packet loss that is caused by severe contention on the radio channel. In Figure 5, the green-line stays in the top of red line, which means that delivering geocast packets using multiforwarding zone (MFZ) with network-coding can have higher average packet throughput. In the proposed multiforwarding zone (MFZ) with network coding scheme, nodes will send the network coded packet with maximum recover rate for packet recovery. As a result, the average packet throughput can be increased.

Figure 6 shows the results of delivering geocast packets using multiforwarding zone (MFZ) mechanisms with 30 nodes within geocasting region. In Figure 6, the source node sends geocast packets to the geocast region. When nodes in the geocast region receive the geocast packets, these nodes then re-broadcast the geocast packets to its neighbors. Blindly flooding induces the broadcast storm, which wastes precious bandwidth by sending redundant packets that will probably collide. The packet throughput increases rapidly at beginning and then drops at time 5 due to packet loss that is caused by severe contention on the radio channel. In Figure 6, the green line stays in the top of red-line, which means that delivering geocast packets using multiforwarding zone (MFZ) with network coding can have higher average packet throughput. In the proposed network-coding based rarest-first packet recovery algorithm with multiforwarding zone (MFZ), nodes will send the network coded packet with the maximum recover rate for packet recovery. As a result, the average packet throughput can be increased. In Figure 6, the red-line (without network coding) is ended at time 17; the green-line (with network coding) is ended at time 19. It means that the multiforwarding zone (MFZ) with network coding continues to send network-coded packets to recover lost packets of its neighbor nodes. Therefore, the simulation time of the green line (with network coding) is longer than that of red line (without network coding).

5. Conclusion

Geocasting can be used to perform the regional broadcast to deliver geographic-related safety, commercials, and advertisements messages. The challenging problem in geocasting is how to deliver packets to all the nodes within the geocast region with high efficiency but low overhead. In this paper,

a network coding based transmission architecture for delivering geocast packets over hybrid sensor-vehicular networks is proposed. By equipping network coding technique for geocast packets according to network status, the proposed algorithm can enhance the utilization of packet transmission and can satisfy the requirement of adaptive telematic services over hybrid sensor-vehicular networks. Simulation results show that the proposed network coding based rarest-first packet recovery algorithm with multiforwarding zone (MFZ) mechanism can have higher average packet throughput.

Acknowledgment

This research is supported by the National Science Council of the Republic of China under Grant NSC 100-2221-E-218-042-.

References

[1] C. Maihofer, "A survey of geocast routing protocols," *IEEE Communications*, vol. 6, no. 2, pp. 32–42, 2004.

[2] Y.-B. Ko and N. H. Vaidya, "Geocasting in mobile ad hoc networks: location-based multicast algorithms," in *Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications*, pp. 101–110, New Orleans, La, USA, 1999.

[3] Y.-B. Ko and N. H. Vaidya, "Location-Aided Routing (LAR) in mobile ad hoc networks," *Wireless Networks*, vol. 6, no. 4, pp. 307–321, 2000.

[4] T. H. Hsu, Y. C. Lo, and M. S. Chiang, "A network coding based geocasting mechanism in vehicle ad hoc networks," in *Proceedings of the International Conference Ubiquitous Computing and Multimedia Applications (UCMA '11)*, pp. 223–232, Republic of Korea, 2011.

[5] I. Stojmenovic, A. P. Ruhil, and D. K. Lobiyal, "Voronoi diagram and convex hull based geocasting and routing in wireless networks," in *Proceedings of the 8th IEEE International Symposium on Computers and Communication*, pp. 51–56, Canada, 2003.

[6] J. Boleng, T. Camp, and V. Tolety, "Mesh-based geocast routing protocols in an ad hoc network," in *Proceedings of the 15th International Parallel and Distributed Processing Symposium*, pp. 1924–1933, San Francisco, Calif, USA, 2001.

[7] Y. B. Ko and N. H. Vaidya, "GeoTORA: a protocol for geocasting in mobile ad hoc networks," in *Proceedings of the International Conference on Network Protocols*, pp. 240–250, 2000.

[8] Y. B. Ko and N. H. Vaidya, "Anycasting-based protocol for geocast service in mobile ad hoc networks," *Computer Networks*, vol. 41, no. 6, pp. 743–760, 2003.

[9] K. Seada and A. Helmy, "Efficient geocasting with perfect delivery in wireless networks," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '04)*, vol. 4, pp. 2551–2556, 2004.

[10] R. Ahlswede, N. Cai, S. Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, 2000.

[11] S. Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Transactions on Information Theory*, vol. 49, no. 2, pp. 371–381, 2003.

[12] D. S. Lun, N. Ratnakar, R. Koetter, M. Médard, E. Ahmed, and H. Lee, "Achieving minimum-cost multicast: a decentralized approach based on network coding," in *Proceedings of the IEEE INFOCOM*, pp. 1608–1617, March 2005.

[13] S. Katti, D. Katabi, W. Hu, and R. Hariharan, "The importance of being opportunistic: practical network coding for wireless environments," in *Proceedings of the 43rd Allerton Conference on Communication, Control, and Computing*, Monticello, Va, USA, 2005.

[14] W. Xiuchao, "Simulate 802.11b Channel within NS2," 2004, http://cir.nus.edu.sg/reactivetcp/report/80211ChannelinNS2_new.pdf.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 307246, 11 pages
doi:10.1155/2012/307246

Research Article

An Energy Efficient Localization-Free Routing Protocol for Underwater Wireless Sensor Networks

Abdul Wahid and Dongkyun Kim

Graduate School of Electrical Engineering & Computer Science, Kyungpook National University, 1370 Sankyuk-Dong, Buk-Gu Daegu, Republic of Korea

Correspondence should be addressed to Dongkyun Kim, dongkyun@knu.ac.kr

Received 4 November 2011; Accepted 1 February 2012

Academic Editor: Tai Hoon Kim

Recently, underwater wireless sensor networks (UWSNs) have attracted much research attention from both academia and industry, in order to explore the vast underwater environment. UWSNs have peculiar characteristics; that is, they have large propagation delay, high error rate, low bandwidth, and limited energy. Therefore, designing network/routing protocols for UWSNs is very challenging. Also, in UWSNs, improving the energy efficiency is one of the most important issues since the replacement of the batteries of underwater sensor nodes is very expensive due to the unpleasant underwater environment. In this paper, we therefore propose an energy efficient routing protocol, named (energy-efficient depth-based routing protocol) EEDBR for UWSNs. EEDBR utilizes the depth of sensor nodes for forwarding data packets. Furthermore, the residual energy of sensor nodes is also taken into account in order to improve the network lifetime. Based on the comprehensive simulation using NS2, we observe that EEDBR contributes to the performance improvements in terms of the network lifetime, energy consumption, and end-to-end delay. A previous version of this paper was accepted in AST-2011 conference.

1. Introduction

The earth is a water planet, where more than 70% of its area spans over water. Only less than 10% of the water volumes (oceans) have been investigated, while a large area still remains unexplored. The exploration of oceans is getting more attention due to their usefulness such as presence of natural resources, defense, and means of Corresponding author: Dongkyun Kim transportation, and so forth. However, the traditional approaches for monitoring oceans have many limitations such as high cost, longer time of accessing the outcome of monitoring, and the unsuitable underwater environment for human presence, and so forth. Hence, underwater wireless sensor networks (UWSNs) are considered as important alternatives for exploring the oceans.

Recently, UWSNs have attracted much research attention from both academia and industry, in order to explore the vast underwater environment. UWSNs enable a large number of applications such as environmental monitoring for scientific exploration, disaster prevention, assisted navigation, and oil/gas spills monitoring, and so forth.

Terrestrial sensor networks (i.e., ground-based sensor networks) are well investigated, and many communication protocols have been proposed for such networks. However, UWSNs have different characteristics than the terrestrial sensor networks. The major difference is the employment of the acoustic signals in UWSNs, in contrast to terrestrial sensor network where the radio signals are used as a communication media. The transition from the radio to the acoustic signals is due to the poor performance of radio signals in water. The radio signals propagate large distances at extra low frequencies, requiring large antennas and high transmission power.

The employment of acoustic signals as the communication media imposes many distinctive challenges on UWSNs. In general, the UWSNs have the following intrinsic characteristics. The acoustic signals have long propagation delay (i.e., 1500 m/sec), that is, five orders of magnitude higher than the radio signals used in terrestrial sensor networks. The available bandwidth is limited due to attenuation and high absorption factor of acoustic signals. The link quality

is severely affected by the multipath fading and refractive properties of sound channels. Therefore, the bit error rates are typically very high [1, 2].

Since the protocols proposed for terrestrial sensor networks are developed on the basis of the radio signals' characteristics such as low propagation delay and high bandwidth. Therefore, they cannot be directly applied to UWSNs. Thus, enormous efforts have been made for designing efficient communication protocols while taking into account the characteristics of the UWSNs.

The protocols proposed for UWSNs have addressed various issues concerning the characteristics of the UWSNs. Particularly, improving the network lifetime is an important issue in UWSNs since the replacement of the batteries of underwater nodes is very expensive due to harsh underwater environment. Therefore, the network protocol in UWSNs should be designed considering the energy efficiency to improve the network life-time. The underwater sensor nodes consume more energy in transmitting than receiving a packet. Therefore, in order to reduce energy consumption, consequently improving network life-time, the number of transmissions needs to be reduced. In addition, one of the important issues for improving the network lifetime is to balance the energy consumption among the sensor nodes. The workload should be equally divided among all the sensor nodes over a path from a source towards a destination.

In this paper, we therefore propose an energy-efficient depth-based routing protocol (named EEDBR) that performs energy balancing and reduces the number of transmissions of sensor nodes in order to improve the network life-time. In EEDBR, while forwarding a data packet from a sensor node to a sink, the packet is transmitted by some selected nodes. The selection of the nodes is based on the depth and residual energy. The process is as follows. Each sender broadcasts the data packet including a list of its neighbors' IDs, which contains only the IDs of the neighbors having smaller depths than the sender. Hence, only the selected neighboring nodes are allowed to forward the packet. Furthermore, EEDBR performs energy balancing by utilizing the residual energy information of sensor nodes. In EEDBR, sensor nodes hold the packet for a certain time before forwarding. The holding time is based on the residual energy of sensor nodes. A node having high residual energy has a short holding time compared to the nodes having low energy. Hence, the node with high residual energy forwards the packet, and the low energy nodes suppress their transmissions upon overhearing the transmission of the same packet. In this way, the energy balancing is achieved. Due to the energy balancing, the sensor nodes consume their energy parallely, and none of the sensor node's battery is exhausted earlier than others. Hence, the overall network life-time is improved.

The rest of the paper is organized as follows. In Section 2, we review some related routing protocols and their problems. In Section 3, our proposed routing protocol, EEDBR, is described in detail. Section 4 presents the performance evaluation of EEDBR. Finally, conclusions are drawn in Section 5.

2. Related Work

In this section, we present some related routing protocols available in the literature. We take into account the well-known routing protocols proposed for UWSNs. We divide this section into two subsections: localization-based routing protocols and localization-free routing protocols.

2.1. Localization-Based Routing Protocols. In this section, some routing protocols which are based on the localization of the sensor nodes are presented. In [3], a vector based routing protocol called (vector based forwarding) VBF was proposed. The data forwarding in VBF is as follows. A source node, having a data packet to transmit, computes a vector from itself towards the destination/sink node. The source node then broadcasts the data packet including its position/location information in the data packet. In VBF, the nodes near the computed vector are used as relay nodes for forwarding the data packet. Among all the receiving nodes of a broadcast from a sender, only the nodes located in a predefined radius around the computed vector participate in the forwarding of the data packet. The employment of the predefined radius allows a reduced number of nodes to forward the data packet. Hence, the proposed scheme employs the concept of controlled flooding in the network. However, the limitation of the proposed scheme lies in requiring the localization of sensor nodes, which itself is a crucial issue in UWSNs. Furthermore, in case of sparse networks, the unavailability of sensor nodes in the predefined radius affects the performance.

In [4], a routing protocol called (hop by hop vector based forwarding) HHVBF was proposed. HHVBF is the successor of VBF, where the vector is computed on per hop basis. In HHVBF, due to the computation of the vector on per hop basis, performance improvements are achieved over VBF. However, HHVBF still requires the localization of sensor nodes, which limits the applicability of the proposed scheme in real environment.

In [5], a routing protocol called FBR (Focused Beam Routing) protocol was proposed. FBR is a cross-layer approach where different transmission power levels are used during the forwarding of the data packet. The sender of the data packet transmits an RTS packet with a certain transmission power level. If a CTS reply is received from a relay node residing closer to the sink node, the data packet is transmitted to that relay node. Otherwise, the transmission power level is increased to a higher level. FBR uses a range of transmission power levels for example, P_1 to P_N. The limitation of the FBR protocol lies in the assumption that the source node knows its own location and the location of the destination/sink node. Furthermore, the use of RTS/CTS during the forwarding of the data packets causes increased delay and excessive energy consumption. In [6], a routing protocol called (directional flooding-based routing) DFR was proposed. DFR is another routing protocol with the assumption of the localization of sensor nodes.

In [7], (sector based routing with destination location prediction) SBR-DLP was proposed. In SBR-DLP, it is assumed that a mobile sink is available in the network and

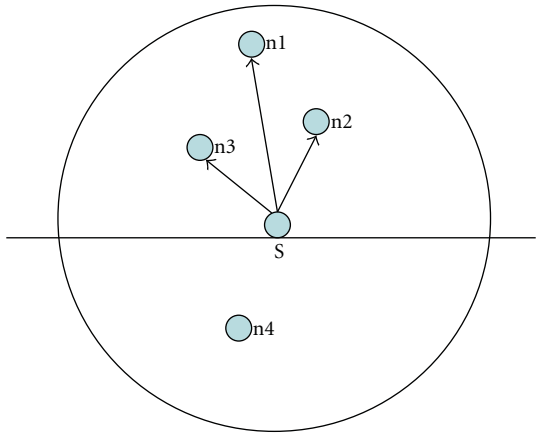


Figure 1: Scenario illustrating drawback of DBR.

that each sensor node is aware of the movement schedule of the mobile sink. The data forwarding process is as follows. A source node broadcasts a chk_ngb packet. The neighboring nodes reply with a chk_ngb_reply packet. The chk_ngb_reply packet contains the sector number of the neighboring nodes. The sectors are computed based on the distance from the vector (i.e., the vector between the source and the sink node). Upon receiving chk_ngb_reply packets from its neighbors, the source node assigns priorities to the neighboring nodes based on the sector number. Then, the neighboring node closest to the mobile sink is selected as a forwarder. The limitations of the proposed scheme are in requiring a localization technique, a large delay due to chk_ngb/chk_ngb_reply packets and the hard assumptions.

2.2. Localization-Free Routing Protocols. In this section, we present some localization-free routing protocols for UWSNs. In [8], a localization-free routing protocol called (depth based routing) DBR was proposed. DBR uses the depth of the sensor nodes as a metric for forwarding data packets. During data forwarding, the sender includes its depth in the data packet. The receiving nodes compare their depths to the depth of the sender. The node having smaller depth participates in forwarding the data packet. Each node has a certain holding time for each data packet, where the nodes having smaller depths have a short holding time compared to the nodes having higher depths. Since only the depth of sensor nodes is used as a metric for forwarding, most of the time, the nodes having smaller depths are involved in forwarding. Hence, such nodes die earlier than the other nodes in the network, which creates the routing holes in the network. Such a scenario is illustrated using Figure 1.

In Figure 1, node S is the sender of the packet, and nodes n1, n2, n3, and n4 are the receiving nodes. According to the approach employed by DBR, nodes n1, n2, and n3 are eligible for forwarding the packet because of having smaller depths than the sending node S. However, every time, the packet is forwarded by node n1 since node n1 has lower depth than the nodes n1 and n2, therefore, having short holding time. Consequently, due to the frequent forwarding, node n1 will die earlier than node n2 and node n3, which will create a routing hole in the network. Since all the nodes have the same approach of forwarding the data packets, routing holes are created all over the network. Due to the routing holes, the network is partitioned into parts which affect the network lifetime.

Furthermore, in DBR, with the increase in network density, the number of redundant transmissions also increases because the probability of small difference among nodes' depths also increases with the network density and the nodes having similar depths also have similar holding times. Due to the long propagation delay in underwater environment, before overhearing the same packet from a sender, a node's timer expires. Consequently, that node also transmits the packet. Hence, all the nodes which are having small differences among their holding times will transmit the packet before overhearing the same packet from other nodes. Thus, a lot of redundant packets will be transmitted, leading to excessive energy consumption.

In [9], H^2-DAB (hop-by-hop dynamic addressing based) routing protocol was proposed. H^2-DAB assigns a unique address (called HopID) to each sensor node based on the hop count from the sink node. The process is as follows. The sink node broadcasts a Hello packet. Each receiving node is assigned a HopID. Then, the receiving nodes increment the HopID and rebroadcast the Hello packet including the updated HopID. Since the HopID is increased hop by hop, the sensor nodes closer to the sink will be assigned smaller HopIDs than the nodes away from the sink. During the forwarding of data packets, the nodes having small HopIDs are selected for forwarding the data packets. Similar to DBR, the nodes having small HopIDs are frequently used for forwarding the data packets. Hence, these nodes having small HopIDs die earlier than the other nodes in the network. In addition, only hop count-based metric is not suitable in a resource-constrained network as UWSN. Furthermore, H^2-DAB uses *inquiry request* and *inquiry reply* packets during the forwarding of the data packets, which is expensive in terms of delay and energy.

In [10], Winston et al. proposed a virtual sink architecture where multiple sinks are assumed connected to each other. In the proposed scheme, each sink broadcasts a Hello packet (called hop count update packet). Upon receiving the Hello packet, each sensor node is assigned a hop count value. During the forwarding of the data packets from a source towards a sink node, these hop count values are used in the selection of a next forwarding node. The proposed scheme's limitations are the redundant transmissions (i.e., the transmission of the same packet towards multiple sinks) and the hard assumption of the connectivity among the sink nodes.

In [11], a network protocol called (multipath power-control transmission) MPT was proposed. MPT uses a cross layer approach, it combines power control with multi-path routing. The proposed scheme is divided into three phases: *multipath routing*, *source-initiated power-control transmission* and *destination node's packet combining.* Initially, the source node transmits a route request packet, the destination node reply with multiple route reply packets. Since the

route request is broadcasted, therefore, the route request packet follows multiple paths towards the destination. Consequently, multiple route reply packets, following different paths, are received by the source node. Then an optimum number of paths are selected based on the path length and the number of paths. Source node also computes optimal energy distribution along a path based on the collected information (i.e., number of paths, number of hops in each path, perhop distance) during path establishment phase. The optimal energy distribution information is also included in the data packet, and based on this information each forwarder selects its transmission power. Finally, upon receiving the data packet, the destination node combines multiple erroneous copies (since some packets might be corrupted) of a packet received from multiple paths into a single copy to recover the original packet. The limitations of the proposed scheme are the redundant packets' transmissions and probabilistic approach used in the computation of optimal energy distribution.

In this paper, similar to DBR, we also employ the depth of sensor nodes for the selection of the forwarding nodes. However, our proposed scheme is different from DBR as follows.

(1) DBR uses only the depth of sensor nodes without taking into account the residual energy of the sensor nodes. In addition, in DBR, there is no method/approach for energy balancing among sensor nodes. In contrast, in our proposed scheme, the energy balancing of sensor nodes is employed in order to improve the network life-time.

(2) In DBR, the number of forwarding nodes increases as the network density increases. However, in our proposed scheme, the number of forwarding nodes is restricted on the basis of not only the depth but also the residual energy of the sensor nodes.

(3) DBR is a receiver-based approach, where the receiving nodes decide whether to forward the received data packet or not. There is a high probability of redundant transmissions in a receiver-based approach due to the lack of neighboring nodes' information such as depth and residual energy. In contrast, our scheme is a sender-based approach where the sender decides the forwarding nodes based on the neighboring nodes' depths and residual energy information. Hence, the sender can select a limited number of suitable forwarding nodes.

3. Our Proposed Protocol: Energy Efficient Depth-Based Routing Protocol (EEDBR)

In this section, we introduce our proposed routing protocol, EEDBR, in detail. EEDBR consists of two phases: knowledge acquisition phase and data forwarding phase. During the knowledge acquisition phase, sensor nodes share their depth and residual energy information among their neighbors. In the data forwarding phase, data packets are transmitted from the sensor nodes to the sink node.

We have divided this section into three subsections: network architecture, knowledge acquisition phase and data forwarding phase.

3.1. Network Architecture. Figure 2 shows the architecture of UWSN. Multiple sink nodes are deployed on the water surface, and the sensor nodes are deployed underwater from the top to the bottom of the deployment region. It is assumed that the sink nodes are equipped with the acoustic and radio modems. These sink nodes use acoustic modems, for communication with the underwater sensor nodes, and the radio modems, for communication with other sinks or an onshore data center. Since the radio communication is much faster than the acoustic communication, the data packet once received at any sink is considered delivered to all sinks and the onshore data center.

3.2. Knowledge Acquisition Phase. During this phase, the sensor nodes share their depth and residual energy information among their neighbors. The purpose of this sharing is to allow the sensor nodes to select the most suitable neighbors as forwarders during the data forwarding phase. When a sensor node has a data packet to send to the sink node, the depth and the residual energy information are used in the selection of forwarding nodes. In this knowledge acquisition phase, knowledge means the depth and residual energy of a sensor node.

The knowledge acquisition process is as follows. Each sensor node broadcasts a *Hello* packet to its one hop neighbors. The *Hello* packet contains the depth and the residual energy of the broadcasting node. The format of the *Hello* packet is shown in Figure 3. Upon receiving the *Hello* packet, the neighboring nodes store the depth and the residual energy information of those sensor nodes having smaller depth. The neighboring nodes only store the information about the sensor nodes having smaller depths since it is obvious that the data packets are transmitted towards the sink nodes residing on the water surface. Hence, storing the depth and residual energy information of all the neighboring nodes is not required, which lessens the burden of storing a large number of data.

It is reported that, in UWSNs, the sensor nodes reside at the same depth. This is because the sensor nodes move with water currents in horizontal direction, and the movements in vertical direction are almost negligible [4]. Hence, the updating of the depth information is not significant. However, the residual energy of the sensor nodes changes over time due to the different operations, that is, transmitting, receiving, processing, and idle listening. Therefore, the residual energy information of the sensor nodes needs to be updated. For this purpose, a distributed approach is employed in our proposed scheme. Each sensor node checks its residual energy on an interval basis. If the difference between the current and previous residual energy of a sensor node is larger than a threshold (i.e., a system parameter), that sensor node broadcasts the *Hello* packet including the updated residual energy to its one-hop neighbors. In this way, the residual energy information of the sensor nodes is updated among the

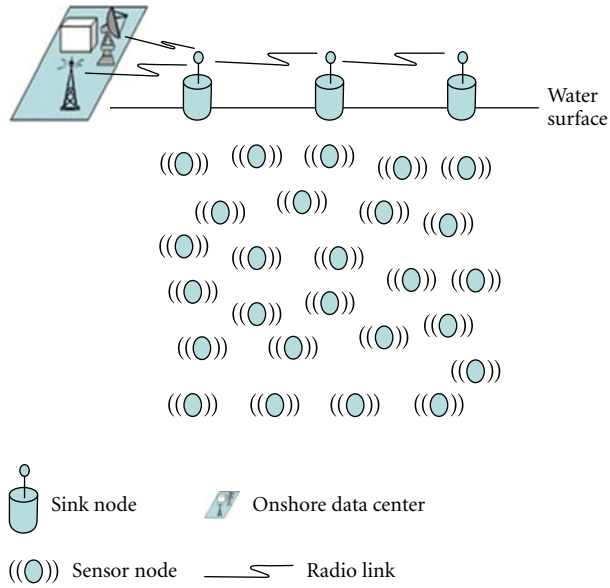


Figure 2: Architecture of an underwater wirless sensor network.

Sender ID	Residual energy	Depth

Figure 3: Format of the hello packet.

neighboring nodes. Furthermore, the knowledge acquisition phase is executed on an interval basis. This is done to update the sensor nodes about their most recent neighboring nodes and their updated residual energy and depths. However, the interval of knowledge acquisition phase is set long in order to avoid the overhead due to the broadcasts of the *Hello* packets. Hence, there is a tradeoff between the overhead and having the updated information about the neighboring nodes.

3.3. Data Forwarding Phase. During this phase, the data packets are forwarded from a source node towards a destination/sink node on the basis of the depth and the residual energy information of the sensor nodes. The information about the depth of the sensor nodes allows the selection of those forwarding nodes which are closer to the sink than the sender of the data packet. In addition, the residual energy information about the sensor nodes is used to select the node having high residual energy among its neighbors. The selection of the node having high energy attempts to balance the energy consumption among the sensor nodes. In EEDBR, since each sensor node has the information about its neighbors' depth and the residual energy, a sending node can select the most suitable next hop forwarding nodes. Therefore, the sending node selects a set of forwarding nodes among its neighbors having smaller depth than itself. The set of forwarding nodes is included as a list of IDs in the data packet.

Upon receiving the data packet, the forwarding nodes hold the packet for a certain time based on their residual energy. The sensor node having more residual energy has a short holding time. The holding time (T) is computed using (1).

$$T = (1 - (\text{current energy}/\text{initial energy})) * \max\_\text{holding\_time} + p, \quad (1)$$

where max_holding_time is a system parameter (i.e., the maximum holding time a node can hold a packet), and p is the priority value.

The priority value is used to prevent multiple forwarding nodes from having the same holding time since the sensor nodes might have the same residual energy level. Therefore, if the holding time is only based on residual energy, the nodes having same residual energy will also have the same holding time. In such a case, the forwarding nodes will forward the packet at the same time. Hence, redundant packets will be transmitted. In order to avoid such redundant transmissions, the priority value is added to the holding time in order to make the difference among the holding times of the forwarding nodes having the same residual energy.

The priority value is computed as follows. The sending node sorts the forwarding list on the basis of the residual energy of the forwarding nodes. Upon receiving the data packet, the forwarding nodes add the priority value to the holding time based on their position in the list. The priority value is initialized with a starting value, and the priority value is doubled with the increase in the position index of the nodes in the list. Hence, due to the different positions in the list, the nodes have different priority values. Consequently, the nodes having the same residual energy will have different holding times even for the same packet.

Figure 4(e) illustrates the scenario where the forwarding nodes have same residual energy, and node S is the sender of

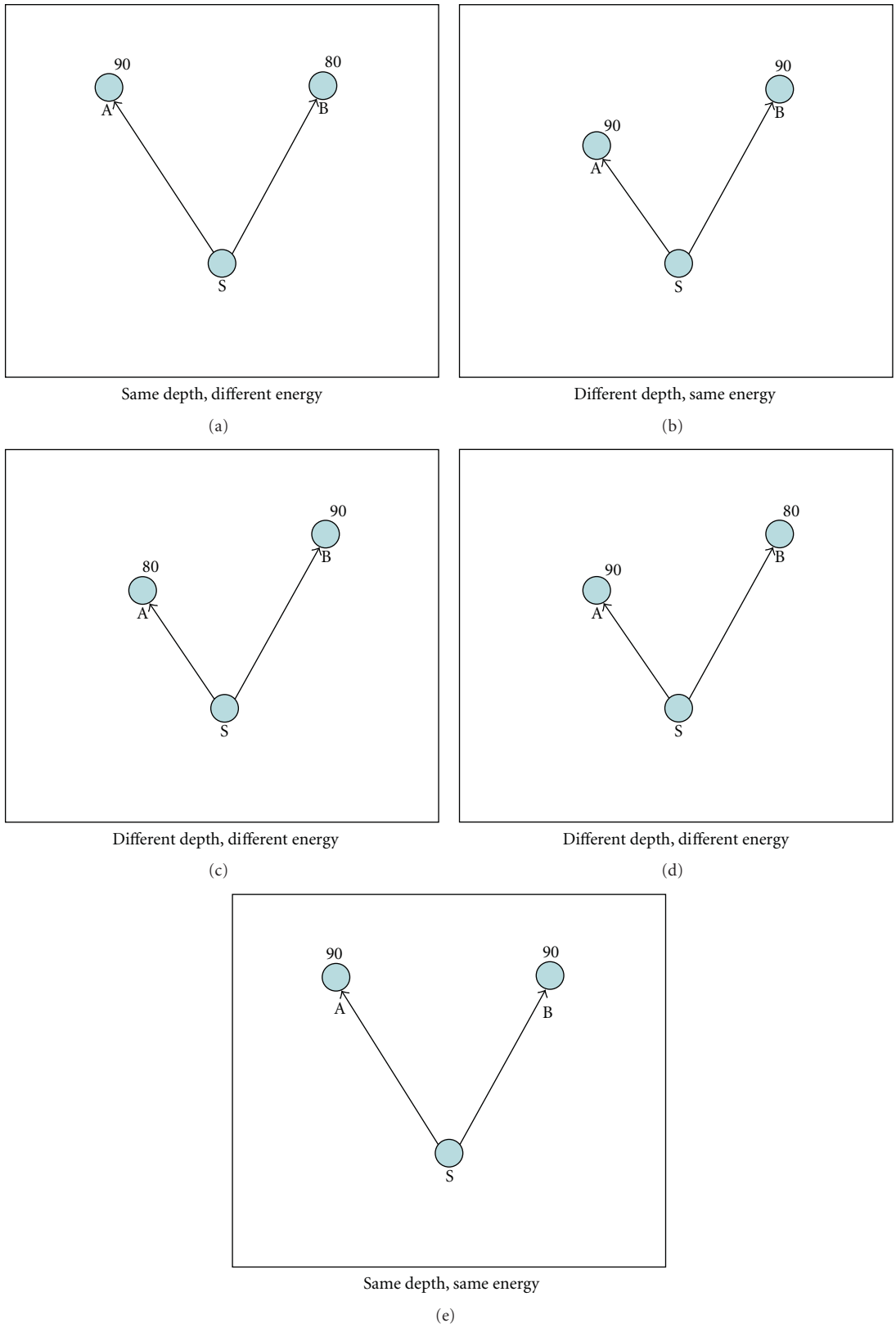


Figure 4: Different possible scenarios during the forwarding of the data packet.

the packet, and node, A and B are the candidate forwarding nodes. The value 90 is assumed as the residual energy of the nodes A and B. As illustrated in Figure 4(e), both nodes A and B have the same residual energy. When these nodes receive the packet, they check their position in the forwarding list. On the basis of the position in the list, both the nodes compute the priority value. Let's assume that the nodes A and B are positioned at the second and third positions in the list and assume that the priority value is started with a starting value 10. Then, the priority value of node A will become 20, and node B will have the priority value of 40, because the priority value is doubled corresponding to the position in the list. Hence, despite of having same residual energy, both the nodes have different holding times for the same packet. Furthermore, since the difference between the holding times of both the nodes is double, a node has an enough holding time for overhearing the same packet from other sensor nodes. In contrast, in DBR, the difference between the holding times of the sensor nodes having similar depths is not long enough for overhearing. Hence, redundant packet transmissions are unavoidable in DBR. In EEDBR, the topmost node in the list has the highest priority because of having the highest residual energy among its neighbors. Therefore, we employ a holding time of zero for the topmost node in the list in order to reduce the end-to-end delay. The topmost node will forward the data packet as soon as it receives the data packet.

During the forwarding of the data packet on the basis of the depth and residual energy, different scenarios are possible as shown in Figure 4 where node S is the sender, nodes A and B are the forwarding nodes, and the values 90 and 80 are assumed as the residual energy values of A and B. Here, we describe how EEDBR responds to such scenarios. In case (a), both the nodes are having same depth. However, the sensor node A forwards the packet since it has more energy than node B. In case (b), both nodes have same residual energy. However, node B forwards the packet because it is located at the lower depth than node A. Similarly, in case (3), node B forwards the packet since it has more energy and also it is located at the lower depth. In case (d), node A forwards the packet because of having more energy. Here, it is also possible to give a priority to the node which has lower depth, which means it is nearer to the sink node. However, for the energy balancing purpose, the node having more energy is preferred. Finally, in case (e), since both the nodes have the same depth and the same residual energy, anyone can be selected for forwarding. As above-mentioned, both the nodes will have different holding times. Hence, one node will transmit the packet, and the other will suppress its transmission upon overhearing the transmission of the same packet.

In UWSNs, the suppressions of packet transmissions contribute to reducing the energy consumption, hence, improving energy efficiency. However, too much suppression of packet transmissions affects the delivery ratio. In some applications such as military surveillance, the delivery ratio is more important than the energy efficiency. Hence, in order to support such applications, we employ an application-based suppression scheme. In our suppression scheme, when the delivery ratio is less than a given delivery ratio threshold, the number of nodes which suppress their packet transmissions is reduced in order to meet the desired delivery ratio. During the forwarding of the data packets, the source includes the number of packets generated by that source. Upon receiving the data packets, the sink node computes the delivery ratio by dividing the number of data packets received at the sink to the number of data packets generated by the source node. If the delivery ratio is less than the desired delivery ratio based on the application requirement, the sink node informs the source node by sending/flooding a packet containing the delivery ratio at the sink. Consequently, the source node includes the delivery ratio value received from the sink into the data packet. Upon receiving the data packet, the forwarding node decides whether to suppress or transmit the packet based on the delivery ratio value in the data packet. Here, a probabilistic approach is used. The forwarding nodes generate a random number. If the random number is less than the delivery ratio value, the packet is transmitted without any suppression even if the same packet is received from other nodes. In this way, the degree of suppressions of packet transmissions is controlled. Hence, there is a tradeoff between the energy efficiency and the delivery ratio, and the proposed scheme, EEDBR, can be switched interchangeably based on the application requirement.

In EEDBR, the data packet's forwarding from a source node to a sink is summarized as follows. Each sender of the data packet includes a list of its neighboring nodes having smaller depths, called forwarding nodes. The list is ordered on the basis of the residual energy values of the forwarding nodes. Upon receiving the data packet, the first node in the list forwards the data packet immediately without waiting. The rest of the forwarding nodes in the list holds the data packet for a certain time computed using (1). During the holding time, upon overhearing the same data packet from another sensor node, the forwarding nodes generate a random number and compare it to the delivery ratio value received in the data packet. The nodes suppress the transmission if the random number is less than the delivery ratio. Otherwise, the data packet is transmitted. In case where no data packet is overheard during the holding time, the data packet is transmitted when the holding time expires. To illustrate further, the operation during the forwarding of the data packet is shown in Figure 5.

4. Performance Evaluation

In this section, we evaluate the performance of our proposed routing protocol, EEDBR, by comparing it to an existing routing protocol in UWSNs called DBR [8]. Since DBR is a representative localization-free routing protocol in UWSNs, we select DBR for the performance comparison.

4.1. Simulation Settings. Simulations were conducted using a commonly used network simulator called NS-2. We performed simulations with a different number of sensor nodes (i.e., 25, 49, 100, and 225). We employed grid and random topologies for the comparisons. In each topology, the transmission range of 250 meters was set for each sensor

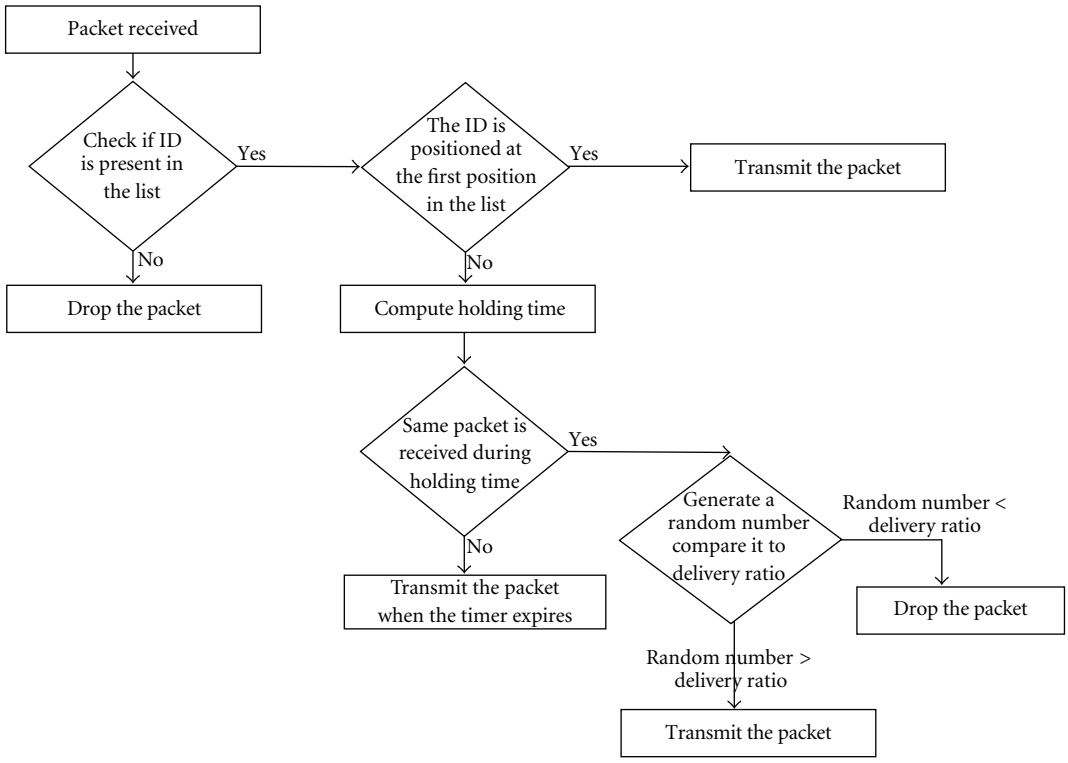


Figure 5: Operation at the forwarding node.

node. The initial energy value of 70 joule was set for all the sensor nodes. Different numbers of sink nodes were used for each topology (i.e., 2, 3, 4, and 6 sink nodes for 25, 49, 100 and 225 sensor nodes). In each topology, two source nodes were randomly selected from the bottom of the deployment region. Each source node generated a data packet of a size of 64 bytes every 15 seconds. The 802.11-DYNAV [12] protocol was used as an underlying MAC protocol. For all topologies, the results were averaged from 30 runs.

4.2. Performance Metrics. We used the following metrics for evaluating the performance of our proposed routing protocol.

Network Lifetime. Network life-time is the time when the first node dies in the network when the energy of that node is fully exhausted.

Energy Consumption. Energy consumption is evaluated through the total amount of energy consumed by the sensor nodes during the forwarding of the data packets from a source towards a destination/sink node.

End-to-End Delay. The end-to-end delay is the time taken by a packet to reach from a source node to a destination/sink node.

Delivery Ratio. Delivery ratio is defined as the ratio of the number of packets successfully received at the sink node to the number of packets transmitted from the source node.

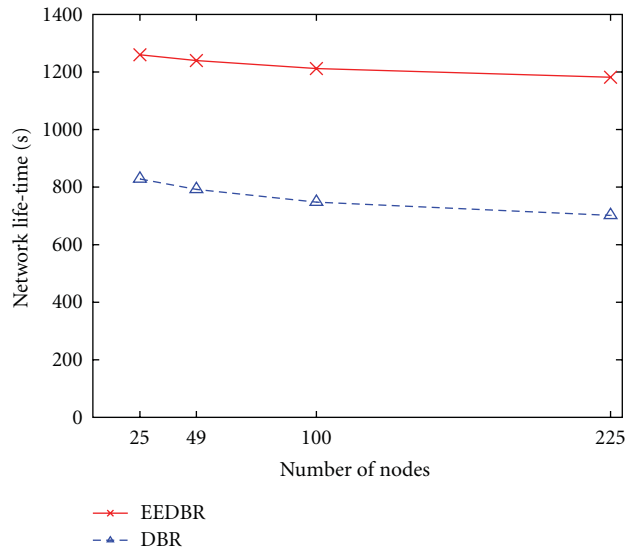


Figure 6: Comparison of network lifetime in random topology.

4.3. Simulation Results and Analysis

4.3.1. Network Lifetime. The network life-time of both the schemes in random and grid topologies is compared as shown in Figures 6 and 10, respectively. EEDBR shows improved performance over DBR. Since DBR selects the nodes having smaller depths to be frequently used for forwarding the data packets. Therefore, the energy of such nodes is exhausted rapidly, and these nodes die very soon. In

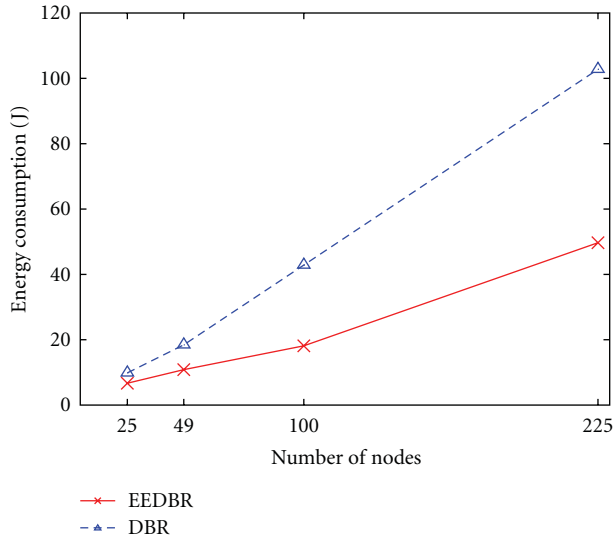


Figure 7: Comparison of energy consumption in random topology.

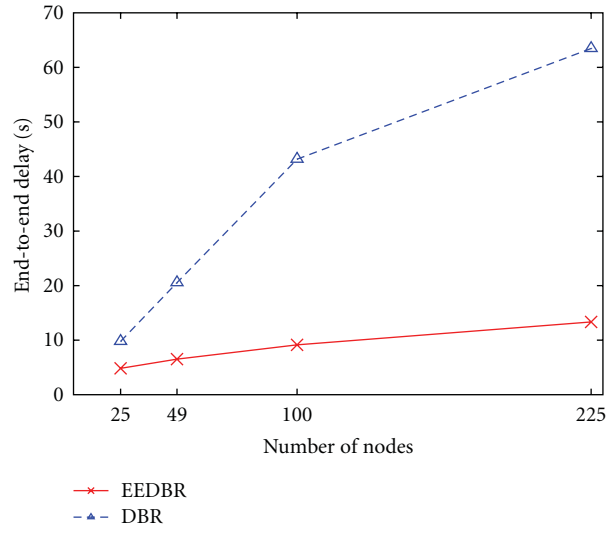


Figure 8: Comparison of end-to-end delay in random topology.

contrast, EEDBR employs the energy balancing among the sensor nodes in order to enable the sensor nodes to consume their energy parallely. Hence, the sensor nodes stay alive for long time. Furthermore, the lower energy consumption is another factor of the improved network life-time in the proposed EEDBR scheme. In EEDBR, a limited number of sensor nodes are allowed to participate in forwarding. The forwarding is not only restricted on the basis of the depth but also the residual energy of the sensor nodes. In addition, DBR cannot avoid redundant packet transmissions. The sensor nodes having similar depths also have similar holding times. Therefore, the same packets are transmitted at the same time in DBR. In contrast, in EEDBR, due to the employment of the priority values, the redundant transmissions of packets do not occur. Hence, the reduction in energy consumptions is achieved, which also improves the battery life-time of the sensor nodes.

4.3.2. Energy Consumption. Figures 7 and 11 shows, the energy consumption of both the schemes in random and grid topologies, respectively. The energy consumption of DBR is higher than the proposed EEDBR protocol due to excessive number of nodes' involvement in forwarding the data packet and redundant packets' transmissions in DBR as mentioned earlier. As shown in the figures, the energy consumption of both the schemes is increasing with the increase in network density. This is because more nodes become eligible for forwarding the data packet with the increase in network density. However, DBR only restricts the number of nodes on the basis of the depth of the sensor nodes. Only utilizing the depth of the sensor nodes can not reduce the number of nodes since sensor nodes have similar depths. In contrast, EEDBR restricts the number of nodes, based on two metrics: the depth and the residual energy. Furthermore, in EEDBR, due to the priority assignment technique, nodes have enough difference in their holding times. Therefore, the nodes holding a packet suppress their transmissions upon overhearing the transmission of the same packet form a high priority sensor node.

4.3.3. End-to-End Delay. The end-to-end delay of both the schemes is investigated as shown in Figures 8 and 12 in random and grid topology, respectively. In DBR, each sensor node holds the packet for a certain time proportional to the depth of the sensor node. Therefore, DBR has a long end-to-end delay. In contrast, EEDBR wants the first node in the list of forwarding nodes to transmit the packet as soon as it receives the packet. Therefore, the delay is reduced only to the propagation delay of the packet. As depicted in the figures in both random and grid topologies, the delay in DBR is continuously increasing with the increase in network density because the number of forwarding nodes also increases with the increase in network density. Since each node holds the packet for a certain time, the overall holding time of the packet also increases. The increase in network density does not affect the end-to-end delay in EEDBR, because each time the first forwarding node in the list has a holding time of zero.

4.3.4. Delivery Ratio. Figures 9 and 13 show the delivery ratio of both the schemes in random and grid topologies, respectively. The delivery ratio is much better in random topology than the grid topology, where the delivery ratio is higher than 94% for both the schemes in random topology. The delivery ratio is more than 90% in grid topology. However, when the network density reaches 225 nodes, the delivery ratio is abruptly dropped to 85%, since the number of collisions also increases with the increase in number of nodes. Relatively, DBR has better delivery ratio than EEDBR. The delivery ratio of DBR is 2 to 3% higher than EEDBR in random topology and 5 to 7% higher in grid topology. This is because DBR makes packets transmitted redundantly where multiple paths are followed to reach the sink node. Hence, the delivery ratio is high in DBR. However, the high

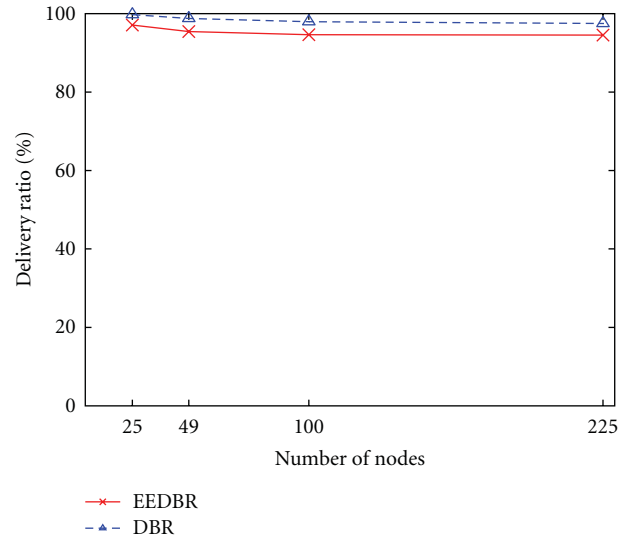


Figure 9: Comparison of delivery ratio in random topology.

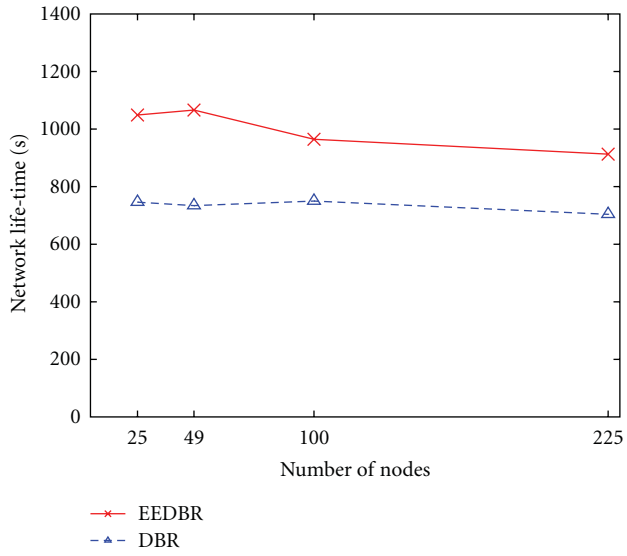


Figure 10: Comparison of network lifetime in grid topology.

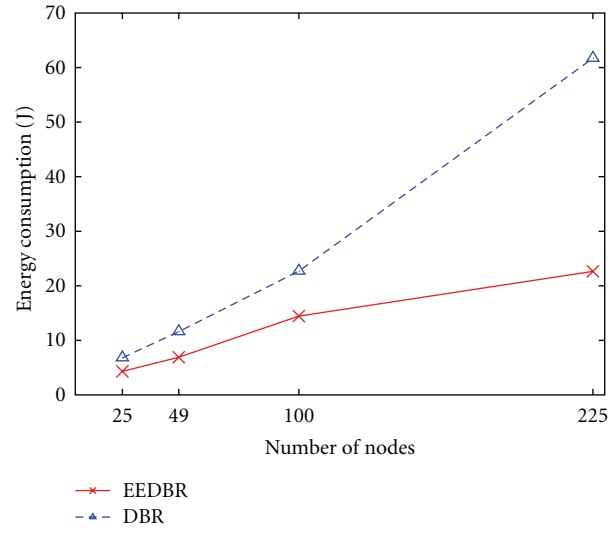


Figure 11: Comparison of energy consumption in grid topology.

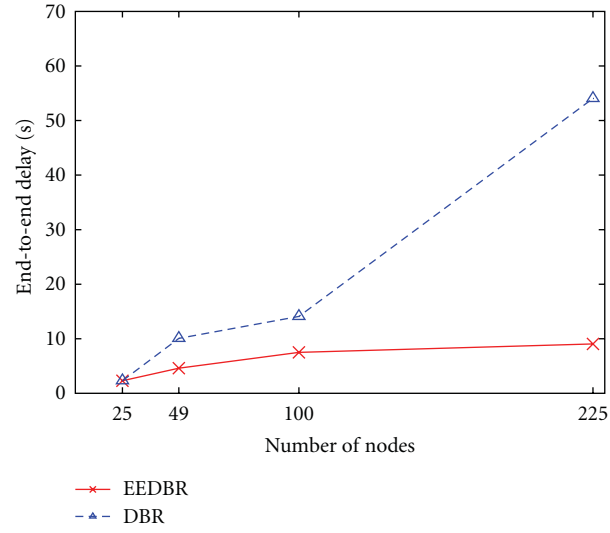


Figure 12: Comparison of end-to-end delay in grid topology.

delivery ratio in DBR is with the expense of excessive energy consumption and increased end-to-end delay.

5. Conclusions

Improving the energy efficiency in underwater wireless sensor networks (UWSNs) is one of the important issues, since the replacement of the batteries of underwater sensor nodes is very expensive due to harsh underwater environment. In this paper, we therefore proposed an energy efficient depth-based routing protocol (named EEDBR) for UWSNs. EEDBR utilizes the depth and the residual energy of sensor nodes as a routing metric. In particular, EEDBR does not require the localization of the sensor nodes which itself is a crucial issue in UWSNs. EEDBR employs a sender-based approach for routing where the sender decides a set of next forwarding nodes in order to reduce redundant transmissions from multiple forwarders. EEDBR has two phases, namely, knowledge acquisition phase and data forwarding phase. In the knowledge acquisition phase, each sensor node shares its depth and residual energy with its neighbors through *Hello* messages. In the data forwarding phase, each sender of the data packet includes a list of its neighboring nodes to the data packet. The set of the neighboring nodes called forwarding set/list is selected based on the depth of the neighboring nodes. Upon receiving the data packet, the forwarding nodes hold the packet for a certain time. The holding time is based on the residual energy of the forwarding nodes. Furthermore, we employed a novel suppression technique for the nodes overhearing the same packet. The degree of suppression of packet transmissions is controlled based on the delivery ratio which is notified by the sink node.

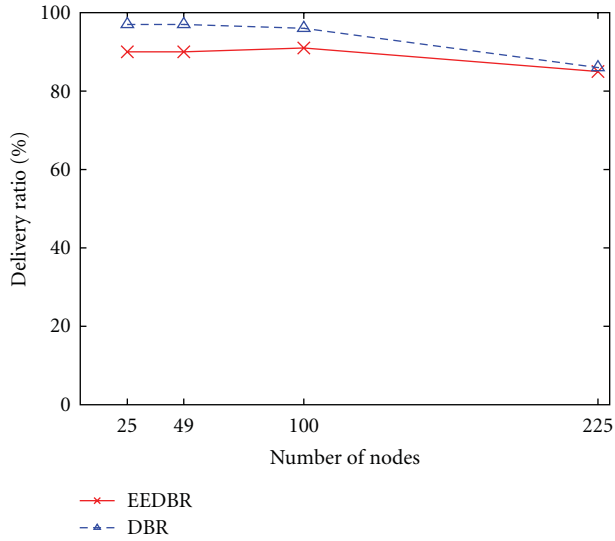


Figure 13: Comparison of delivery ratio in grid topology.

Through NS-2 network simulations, the EEDBR protocol was compared to a representative routing protocol in UWSNs called DBR [8]. Based on the comprehensive simulation, we observed that EEDBR contributes to the performance improvements in terms of network lifetime, energy consumption and end-to-end delay, while keeping the delivery ratio almost similar to the compared routing protocol.

Acknowledgment

This work was supported by Defense Acquisition Program Administration and Agency for Defense Development under the contract UD100002KD. This paper is an extension of our conference paper appeared in AST-2011 [13].

References

[1] J. Heidemann, W. Ye, J. Wills, A. Syed, and Y. Li, "Research challenges and applications for underwater sensor networking," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '06)*, pp. 228–235, Las Vegas, Nev, USA, April 2006.

[2] J. Llor, E. Torres, P. Garrido, and M. Malumbres, "Analyzing the behavior of acoustic link models in underwater wireless sensor networks," in *Proceedings of the 4th ACM Workshop on Performance Monitoring, Measurement and Evaluation of Heterogeneous Wireless and Wired Networks (PM2HW2N '09)*, pp. 9–16, Canary Island, Spain, 2009.

[3] P. Xie, J.-H. Cui, and L. Lao, "VBF: vector-based forwarding protocol for underwater sensor networks," in *Proceedings of the IFIP Networking*, pp. 1216–1221, Lisbon, Portugal, 2006.

[4] N. Nicolaou, A. See, P. Xie, J. H. Cui, and D. Maggiorini, "Improving the robustness of location-based routing for underwater sensor networks," in *Proceedings of the Mts/Ieee Kobe-Techno-Ocean (OCEANS '07)*, pp. 1–6, June 2007.

[5] J. Jornet, M. Stojanovic, and M. Zorzi, "Focused beam routing protocol for underwater acoustic networks," in *Proceedings of the 3rd International Workshop on Underwater Networks (WUWNet '08)*, pp. 75–81, September 2008.

[6] D. Hwang and D. Kim, "DFR: directional flooding-based routing protocol for underwater sensor networks," in *Proceedings of the Mts/Ieee Kobe-Techno-Ocean (OCEANS '08)*, pp. 1–7, Quebec City, Canada, September 2008.

[7] N. Chirdchoo, W.-S. Soh, K.-C. Chua et al., "Sector-based routing with destination location prediction for underwater mobile networks," in *Proceedings of the IEEE International Conference on Advanced Information Networking and Applications (WAINA '09)*, pp. 1148–1153, 2009.

[8] H. Yan, Z. Shi, and J. Cui, "DBR: depth-based routing for underwater sensor networks," in *Proceedings of the IFIP (Networking '08)*, pp. 16–1221, Singapore, 2008.

[9] M. Ayaz and A. Abdullah, "Hop-by-hop dynamic addressing based (H2-DAB) routing protocol for underwater wireless sensor networks," in *Proceedings of the International Conference on Information and Multimedia Technology (ICIMT '09)*, pp. 436–441, Jeju Island, Korea, December 2009.

[10] W. G. Seah and H.-X. Tan, "Multipath virtual sink architecture for underwater sensor networks," in *Proceedings of the Mts/Ieee Kobe-Techno-Ocean (OCEANS '06)*, Singapore, May 2007.

[11] Z. Zhou and J.-H. Cui, "Energy efficient multi-path communication for time-critical applications in underwater sensor networks," in *9th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc '08)*, pp. 221–230, Hong Kong, China, May 2008.

[12] D. Shin and D. Kim, "A dynamic NAV determination protocol in 802.11 based underwater networks," in *Proceedings of the IEEE International Symposium on Wireless Communication Systems (ISWCS '08)*, pp. 401–405, Reykjavik, Iceland, October 2008.

[13] http://www.sersc.org/AST2011/.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 708762, 9 pages
doi:10.1155/2012/708762

Research Article

An Optimization Scheme for M2M-Based Patient Monitoring in Ubiquitous Healthcare Domain

Sungmo Jung,[1] Jae Young Ahn,[2] Dae-Joon Hwang,[3] and Seoksoo Kim[1]

[1] *Department of Multimedia, Hannam University, Daejeon 306-791, Republic of Korea*
[2] *Standards Research Center, Electronics and Telecommunications Research Institute, Daejeon 305-700, Republic of Korea*
[3] *Department of Information and Communication Engineering, Sungkyunkwan University, Suwon 440-746, Republic of Korea*

Correspondence should be addressed to Seoksoo Kim, sskim0123@naver.com

Received 3 November 2011; Accepted 30 January 2012

Academic Editor: Tai Hoon Kim

In ubiquitous healthcare systems, machine-to-machine (M2M) communication promises large opportunities as it utilizes rapidly developing technologies of large-scale networking of devices for patient monitoring without dependence on human interaction. With the emergence of wireless multimedia sensor networks (WMSNs), M2M communications improve continuous monitoring and transmission and retrieval of multimedia content such as video and audio streams, images, and sensor data from the patient being monitored. This research deploys WMSN for continuous monitoring of target patients and reports tracking for preventive ubiquitous healthcare. This study performs optimization scheme movement coordination technique and data routing within the monitored area. A movement tracking algorithm is proposed for better patient tracking techniques and aids in optimal deployment of wireless sensor networks. Results show that our optimization scheme is capable of providing scalable and reliable patient monitoring results.

1. Introduction

The rapid increase in the size of aging population combined with the rise in the healthcare costs is demanding cost-effective and ubiquitous patient monitoring systems. This challenge can be addressed by a reliable patient monitoring solutions for both short-term home healthcare and long-term nursing home care for stationary and mobile patients. A number of these devices communicating through wireless technologies can form a wireless body area network (WBAN) and consist of a set of mobile and compact intercommunicating sensors either wearable or implanted into the human body which provides a new enabling technology for patient monitoring. The emergence in wireless multimedia sensor networks (WMSNs) enables continuous monitoring and transmission and retrieval of multimedia content such as video and audio streams, images, and sensor data from the patient being monitored from remote locations.We made the highlighted changes in the second and third addresses as per official websites.

The growing interest in sensor applications has created a need for protocols and algorithms for large-scale self-organizing ad hoc networks, consisting of hundreds or thousands of nodes. Hence, in the past decade, wireless sensor networks (WSNs) have been the topic of considerable research effort due to their potential for healthcare applications and their ability of being incorporated in M2M networks. Machine-to-machine (M2M) communication is a promising technology in healthcare due to short range wireless networking, wireless mobile networks, and advances in device networking [1]. Although M2M networks do not only consist of sensors, WSNs are a key component of M2M communication, thus, they are referred to as M2M networks [2]. WMSNs are not only capable of sensing, controlling, and actuating scalar data but also capable of sensing and controlling multimedia contents. The network nodes are generally equipped with data processing and communication capabilities which are used for collecting and disseminating healthcare data.

M2M technology is capable of building wireless M2M ecosystems covering a wide range of healthcare applications. With increased processing power, it would enable to jointly deliver federated healthcare services to users that fully leverage the power of M2M technology. With its capability of capturing and analyzing the massive amount of data available in all kinds of smart devices, M2M is a business concept used for automatic transmission of various data from remote sources by wired, wireless, radio, and other transmission technologies.

In monitoring applications, WSNs are modeled as graphs for routing and coverage of sensor devices. The coverage of a sensor network represents the quality of monitoring that the network can provide, for instance, how well an area of interest is monitored by ubiquitous sensors and how effectively a wireless sensor network can locate and monitor patients. Wireless sensor networks can assist in detecting target patient as well as keep the movement information of the patient. Sensor nodes establish face structure to track the designated target patient.

This research studies the coverage of the wireless multimedia sensor network based on the dynamic aspect of the network that is dependent on the movement of wireless sensors. Specifically, we are interested in the coverage resulting from the mobility of ubiquitous sensors for mobile patient monitoring. We represent the performance criteria as a parametric mathematical function of the distributed wireless sensor positions and perform a numerical optimization procedure on the defined function. In this optimization scheme, we limit our current focus to problems of detectability, that is, the system's design goal is to find mobile targets that are moving inside a monitoring area. For the goal of optimization, we optimize sensor placements with the goal of maximizing the probability of successful target tracking for a set of wireless sensors. Additionally, we study the effect of node mobility, fairness across multiple simultaneous paths, and patterns of packet loss, confirming the system's ability to maintain stable routes despite variations in node location and data rate.

2. Related Works

The emergence of low-power, single-chip radios based on the 802.15.4 [3] standards has precipitated the design of small, wearable, truly networked medical sensors. Several design issues and techniques for WSNs describing the physical constraints on sensor nodes, applications, architectural characteristics, and the communications protocols proposed in all layers of the network stack have been addressed by [4, 5]. The use of wireless sensors in invasive and continuous health-monitoring systems was presented by [6]. An implementation of bedside patient monitoring was developed by [7], while [8] implemented a WAP-based telemedicine system. A comprehensive list of recently proposed routing protocols is presented by [9], and routing algorithms used in WSNs were classified as data-centric, hierarchical, and location-based. The early literature on wireless networking addressed the design of efficient routing algorithms without optimization of the energy required to send the messages. Additionally, a comprehensive survey of routing techniques proposed for wireless sensor networks is also presented by [10]. The techniques addressed routing challenges and design issues that may affect the performance of routing protocols in WSNs. The growing interest in sensor applications has created a need for protocols and algorithms for large-scale self-organizing ad hoc networks, consisting of hundreds or thousands of nodes. Although M2M networks do not only consist of sensors, wireless sensor networks (WSNs) are key components of M2M communication that sometimes sensor networks are referred to as M2M networks [11]. Despite the keen interest in M2M and great value in building such a system, M2M is still relatively new and the technology faces several significant challenges.

The coverage of a wireless sensor network represents the quality of monitoring that the network can provide, for instance, how well an area of interest is monitored by wireless sensors and how effectively a sensor network can detect target patients. While the coverage of a sensor network with immobile sensors has been extensively explored and studied by [12, 13], researchers have recently studied the coverage of mobile sensor networks. Most of this work focuses on algorithms for repositioning of sensors in desired positions in order to enhance monitoring and tracking of the network coverage [14–16].

3. Ubiquitous Healthcare Design Requirements

Typically, the requirements for a ubiquitous sensor network design depend heavily on the specific application and deployment environment. In this chapter, we identify several characteristics that nearly all ubiquitous sensor networks would share.

(i) Mobility of devices: both patients and healthcare are mobile, requiring that the communication layer adapts rapidly to changes in link quality. For example, if a multihop routing protocol is in use, it should quickly find new routes when a doctor moves from room to room during rounds.

(ii) Platforms for wearable sensor: healthcare applications generally require very small, lightweight, and wearable sensors. Existing mote platforms are good for demonstrations, but we have found that the large battery packs and protruding antennas are suboptimal for delivery of medical services.

(iii) Multiple receivers: we expect that the data from a given patient will typically be received by multiple doctors or healthcare personnel caring for the patient. This suggests that the network layer should support multicast semantics.

(iv) Communication reliability: in healthcare domains, a great emphasis is placed on data availability. Although intermittent packet loss due to interference may be acceptable, persistent packet loss due to congestion or node mobility would be problematic. Depending on the sensors in use, sampling rates

may range anywhere from less than 1 Hz to 1000 Hz or more, placing heavy demands on the wireless channel.

3.1. M2M Communication. The design of the ubiquitous healthcare system is based on M2M technology. M2M is a combination of various heterogeneous electronic, communication, and software technologies. A typical M2M system comprises the following basic components: intelligent sensor devices, M2M area network, M2M gateway, communication network, and remote client or application [17]. In ubiquitous healthcare system, intelligent devices include wireless multimedia sensors, actuators, RFID tags, wireless body sensors, mobile devices, PC or workstation that incorporates a communications among them.

As described above, the M2M gateway is responsible for extracting raw data from an intelligent device and preparing it for the network. The gateway uses a protocol or driver to interact with the intelligent device and translate the data into a format that another device, application, or human can understand. Mainly, an M2M gateway facilitates communication among the various devices and provides a connection to a backhaul that reaches the Internet. With Internet serving as communications network in an M2M application, it is the central connection component between an intelligent device and a remote client. It provides communications between the M2M gateways and the patients being monitored. The server is the destination of the information.

3.2. System Architecture. The design and deployment of these wireless sensor networks can be a cost-effective alternative to the growing number of sensor networks. In this paper, we illustrate a typical scenario in a home for the aged where a patient is monitored by a caregiver or a medical staff regularly. Consider an elderly patient who has a systemic, arterial hypertension and needs to check his blood pressure from time to time. One solution is to keep his blood pressure under control. This can be done by continuously monitoring and logging his vital parameters. If he is having an emergency situation while being alone in a room, the emergency help may not be available immediately. This situation can be improved by doing patient monitoring using wireless sensor networks. This will enable monitoring for mobile and stationary patients in indoor and outdoor environments.

The development of WMSN allows real-time analysis of sensors' data, provides guidance and feedback to the user, and generates warnings based on the user's state, level of activity, and environmental conditions related to patients. WMSNs include a number of wireless multimedia sensors to generate necessary patient information which includes blood pressure, heart rate, temperature, ECG, EKG, and brain-related information. Additional information is also measured and monitored such as video, audio, current location, motor activity, and other relevant data. The system architecture of a WMSN is shown in Figure 1 where it is composed of a set of wireless sensors attached to the body.

For M2M communication, the ubiquitous healthcare system is based on a publish/subscribe routing framework, allowing multiple sensor devices to relay data to all receivers that have registered an interest in that data. This communication model fits naturally with the needs of medical applications where a number of caregivers may be interested in sensor data from overlapping groups of patients. A discovery protocol is provided to allow end-user devices to determine which sensors are deployed in the network, while a query interface allows a receiving device to request data from specific sensors based on type or physical node address. The query interface also provides a filter facility, whereby a query can specify a simple predicate on sensor data that will transmit only when the data passes the filter. For example, a doctor might request data on a patient only when the vital signs fall outside of a normal range.

4. System Assumptions

In this research, we show our assumptions on the distributed wireless sensor network and target models in target track parameter scenarios. Our goal is to study the coverage of wireless sensor networks with regards to patient tracking and monitoring and obtain the estimation models with respect to the distributed wireless sensors' computation and measurements. We consider patient monitoring systems where multiple sensor detections must occur over a given time interval. Such scenario occurs where data transmission is taking place between sensors. The dynamic aspect of the network coverage depends on the movement of sensors in the network which can be stationary or mobile where patients are moving randomly. As such, this study focuses on a bounded area such as hospital where patients are confined in a predefined area of monitoring. Additionally, a sensor can detect the accurate location of the patient, because the sensor utilizes trilateration to compute the object's location. The trilateration has been proposed in [18]. A sensor node knows its location, and this information can be acquired from global positioning system (GPS) or other mechanisms.

We consider a patient monitoring region $Z \subset \mathbb{R}^2$ with a radius r. A wireless sensor can sense the patient and the environment and detect events within its sensing area which is represented by a disk with radius r centered at the sensor. Within Z, the finite set of wireless sensors is assumed to have identical functionalities. In general, the functionality of individual sensors is defined by a radius of tracking $R(Z)$ and the associated probability of detection $P(Z)$ such that any point within the monitoring region is tracked with probability $P(Z)$. We assume that n represents the number of sensors deployed that track patient located at random position during specified time t. We also assume that a single patient is present and moving with speed v in the sensor region at time interval $[0, t]$.

The monitoring region represents the set of all possible sensor locations which can track the patient in an uneven velocity. The monitoring region is defined as a function of tracking position $p_{\mathrm{TP}} \in R^d$, tracking direction θ_{TH} relative to its tracking origin, and the tracking distance d_{TD} that the target patient travels during the time interval. We assume that each target patient moves independently of each

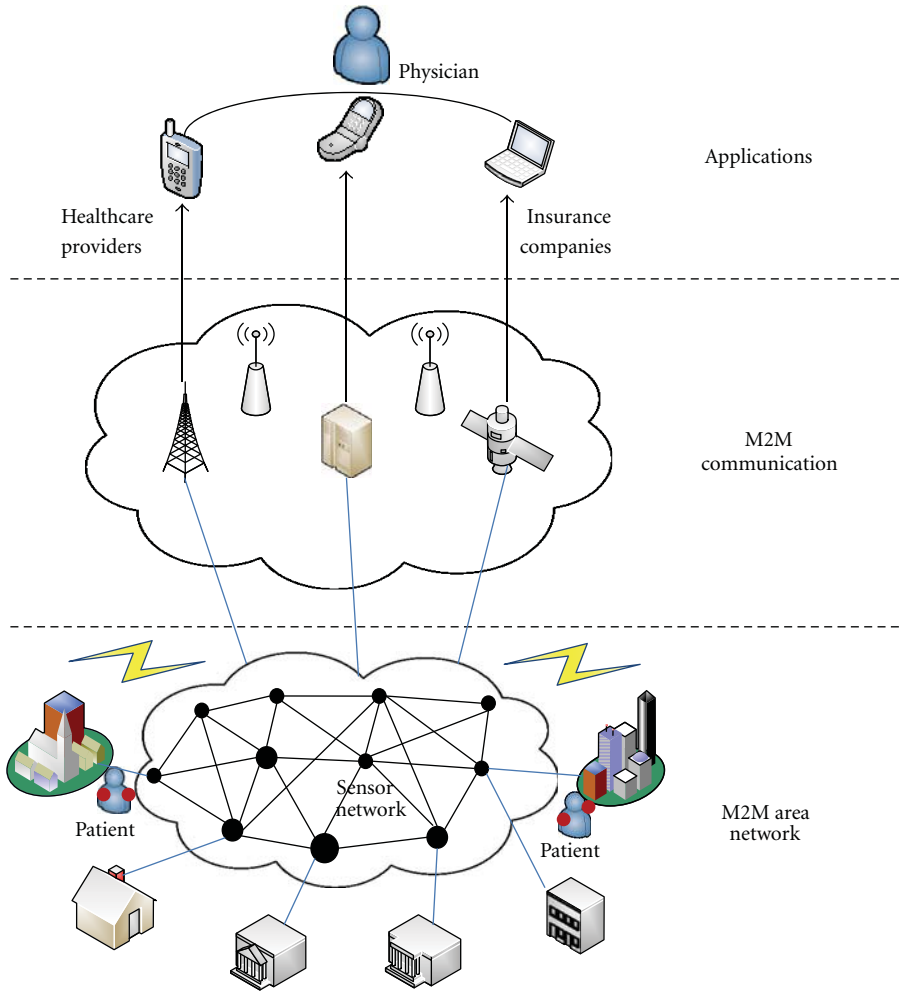


Figure 1: WMSN for patient monitoring.

other and with coordination among them. The number of wireless sensors located in monitoring region Z, $N(Z)$, follows a Poisson distribution of parameter $\lambda\|A_z\|$ where $\|A_z\|$ represents the area of the monitoring region given by

$$P_z(N(Z) = k) = \exp^{-\lambda A_z} \frac{(\lambda A_z)k}{k!}, \quad (1)$$

where λ is the Poisson process parameter. Since each sensor covers a monitoring region with a radius r, the configuration of the wireless sensor network can be initially described by a Poisson probability model $G(\lambda, r)$. Sensors in a stationary sensor networks stay in place after being deployed and network coverage remains the same as that of their initial configuration while, in a mobile sensor network depending on the mobile platform and application scenario, sensors can choose from a wide variety of mobility strategies, from passive movement to highly coordinated and complicated movement. For wireless sensors, the area coverage of a wireless sensor, at any given time instant $t > 0$ time, relative to the monitoring region Z is defined as

$$\gamma_z(t) = 1 - \exp^{-\lambda\pi r^2}, \quad (2)$$

where γ is the probability that a single patient is present in a monitoring region. The localization of patients in the monitoring region can be solved as a nonrandom parameter estimation problem as follows. Let $p_j \in R^d$, $j \in \{1, \ldots, n\}$, which denotes the position of N_Z sensors in a monitoring region $Z \subseteq R^d$, and let $q_0 \in Z$ be the unknown track position to be estimated by means of the movement measurement model:

$$\chi_j(q) = \varphi\left(\left\|q - p_j\right\|\right), \quad q \in Q, \quad (3)$$

for $j \in \{1, \ldots, n\}$. The stacked vector of measurements at a given instant is a random vector normally distributed as

$$Z = \begin{bmatrix} \chi_1 \\ \vdots \\ \chi_n \end{bmatrix} \sim N\left(\begin{bmatrix} \varphi(\|q - p_1\|) \\ \vdots \\ \varphi(\|q - p_n\|) \end{bmatrix}, R\right), \quad (4)$$

where $R > 0$ as the $N \times N$ covariance matrix. In here, we consider the target patient with assumed position z_{TP} moving in direction θ_{TH} and speed v. We make the assumptions that each sensor moves in discrete time along the bounded

region and its sensors detect its immediate clockwise or counterclockwise neighbors and acquire the corresponding distances. Figure 2 shows sensor movement along the boundary of the monitoring region with respect to point q.

Additionally, we define the probability of a sensor being within monitoring region Z and tracking the target patient as P_{TP}. For a distributed tracking approach, we require at least k sensors to be within the region Z and to track the target patient independently, for the particular track patient associated with Z to be tracked and monitored as shown in Figure 3.

While a target patient is tracked, the sensor network has to record the target tracks. A source S obtains the target location that is informed from the first target node p after completing a target discovery process, and then S starts to move toward the first target node p's location. When S reaches the position of the first target node p, S queries the target node for next position. The target node p informs S the target's or next target node's location information. If the target is still located, the source S moves to the target location and catches the target p. If the target has left, the target node n informs S next node q's location. Then, the source moves toward the next target node again. The node p also informs the next node q the information that the source S will reach q. This node does not need to track target for source S anymore. The next target node q becomes the first node. This process is repeated until the source catches target.

Hence, we require k out of N sensors to track the target within Z with equal probability $P_{D\lambda}$. This is represented as binomial probability distribution written as

$$P_{TZ}(N_Z = k) = \binom{N}{k} \left(P_{D\gamma}\right)^k (1 - P_{D\gamma})^{N-k}, \tag{5}$$

where P_{TZ} is the probability of tracking a target patient using the distributed detection criteria. There are cases where it is hard to approximate the presence of large number of sensors or a smaller area covered by a specific sensor. To do this, we provide approximation of a large number of sensors N_Z and small individual sensor coverage as defined by

$$P_{TZ}(N_Z = k) = \exp\left(-NP_{D\gamma}\right) \sum_{m=0}^{k} \frac{\left(NP_{D\gamma}\right)^m}{m!}, \tag{6}$$

where we converge the binomial probability distribution to a Poisson probability distribution to approximate a large number of sensors. In order to optimize the sensor density function $f(z)$, it is convenient to represent the density in a parameterized form. This optimization approach is Fisher Information Matrix (FIM) [19]. Here, the sensor area coverage relative to its movement at a specified time t is represented by a sum of weighted curves of Gaussian mixtures as represented by

$$\gamma_Z(t) \frac{1}{2\pi\sigma^2} \exp\left(-\frac{1}{2\sigma^2}\left(p - q_j\right)^T \left(p - q_j\right)\right). \tag{7}$$

These Gaussian measures are well suited to represent unknown smooth functions. Our implementation was limited to approximating the reasonable number of mixture terms to O(55).

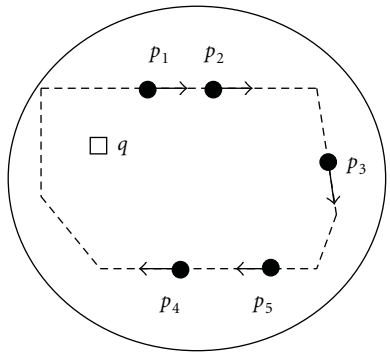


Figure 2: Sensor movement coordination.

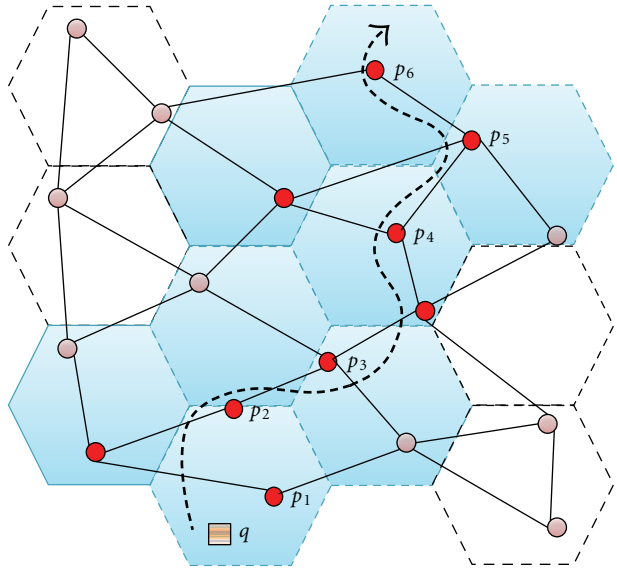


Figure 3: Movement tracking.

4.1. Routing Mechanism. This healthcare system is based on a publish and subscribe routing framework in which sensors publish relevant data to a specific channel and end-user devices subscribe to channels of interest. Publish and subscribe communication decouples the concerns of devices generating data from those receiving and processing it. Practical implementation of a publish/subscribe model must take a number of considerations into account. First, wireless multimedia sensors should not publish data at an arbitrary rate, since the wireless channel has limited bandwidth. Second, given that publishers and subscribers are not necessarily within access range, some form of multihop routing is necessary. Third, the communication layer should take mobility into account when establishing routing paths. In the healthcare scenario, patients and healthcare personnel are mobile. Many patients may be ambulatory and free to roam around the house or in the building.

A good energy-aware routing technique should balance two different goals: choosing a path with maximal residual energy and choosing a path with minimal energy consumption. Of various existing protocols, routing layer protocol used in this healthcare system is based on the adaptive

demand-driven multicast routing (ADMR) protocol [20]. ADMR is chosen due to its simplicity and extensively application in simulation. The publish and subscribe commands allow a node to state that it wishes to associate with a particular channel, while leave terminates a publish and subscribe request. ADMR establishes multicast routes by assigning nodes to be forwarders for a particular channel. A forwarder simply rebroadcasts any messages that it receives on a given channel, using duplicate suppression to avoid multiple transmissions. Nodes are assigned as forwarders through a route discovery process that is initiated when a patient device requests to publish data. Multicast routing allows nodes to avoid transmitting redundant data; for example, if multiple doctors subscribe to vital signs from the same patient, the patient need only transmit its data once to the channel, where it will be forwarded to each recipient.

4.2. Discovery Protocol. In order for the wireless sensor nodes to discover each other and determine the capabilities of each sensor device, a simple discovery protocol is layered on top of the ADMR framework. ADMR supports a special-case broadcast channel that uses a simple controlled flooding mechanism to deliver a message unreliably to every node in the network. Each wireless sensor node periodically publishes metadata about itself, including node ID and sensor types that it supports, to the broadcast channel. Receiving devices that wish to learn about other nodes in the network can subscribe to the broadcast channel to receive this information. Note that the metadata information about a node is static and is not updated frequently. It would be straightforward to reduce the number of broadcast messages by performing in-network aggregation of this metadata.

4.3. Movement Tracking Algorithm. This section presents the algorithm for patient tracking. The goal of this algorithm is the decentralized movement coordination of wireless sensors and localization of target patients. This algorithm assumes a constant $\kappa \in [0, 1/2]$ and information of the target position q. The algorithm is presented in Algorithm 1.

5. Optimization Scheme

This section presents the algorithm for patient tracking. The goal of this algorithm is the decentralized movement coordination of wireless sensors and localization of target patients. This algorithm assumes a constant $\kappa = [0, 1/2]$ and information of the target position q. The algorithm is presented below.

In this section, we will present optimization scheme to compute the area coverage relative to its movement of the wireless sensors. In order to optimize the area coverage of movement coordination of wireless sensors, we require an efficient approach to numerically evaluate the multidimensional integral. As described above. The optimization goal is to find the area coverage which results in the maximum of the probability P_{TZ}, where the function P_{TZ} depends on sensor positions parametrically through the highly nonlinear function $\gamma_z(t)$ which is parameterized by a Gaussian mixture. Hence, the performance measure P_{TZ} is effectively parameterized by the Gaussian weights w_j. According to the general optimal control problem formulation in [21], our optimal mobile sensor area coverage relative to its movement can be formulated as follows.

Maximize

$$\gamma_z(t) P_{TZ} \tag{8}$$

subject to the following constraints

$$\sum_{j=1}^{N} w_j = 1, \quad w_j \geq 0 \quad \forall j. \tag{9}$$

The representation of the area coverage relative to its movement $\gamma_z(t)$ is a mixture of circular Gaussian components defined with fixed position and covariance parameters and variable weights w_j. Heuristics are implemented to determine the number and variance of the components in the mixture for performance optimization. The number and variance of the components in the mixture also depend on the scaling of the search region relative to the sensor parameters. Hence, the objective function based on the assumptions is dependent on the sensor coverage area relative to through the defined weight parameters.

6. Evaluation

This section presents the evaluation of the ubiquitous healthcare system utilizing wireless multimedia sensor networks such as video recorder and audio sensor placed inside the house. Although the location of each node is fixed, this testbed affords us the opportunity to measure communication reliability and throughput under a wide range of link conditions and data rates. Also, wireless body sensors are attached to patients as well as mobile devices to aid the transmission of patient information. The system enables forwarding of messages to and from sensor device for the control and monitoring of the patient and the environment.

The setup enables to run tests with many different parameters without having to reprogram the sensor devices each time. In each experiment, we experiment wireless sensors on each patient device that generates data at a constant rate. Each experiment was executed for at least 2 minutes, and statistics were calculated after removing the first 60 seconds of each trace to avoid measuring startup effects.

This experiment measures three separate sender-receiver pairs with different number of radio hops in the ADMR path. Increasing the transmission rate leads to degradation in reception rate due to dropped packets issuing queries, receiving data, retrieving statistics, and so forth.

Figure 4 shows the packet reception ratio (the number of received packets divided by the number of transmitted packets) for three separate sender-receiver pairs. In all three cases, the same node is used as the sender, while the receiving node is varied. Receivers were selected to vary the number of radio hops along the ADMR path. Note that the hop count

Set time to t
While sensor agent $i = 1$ to n do
(1) Get the estimate position from central server.
(2) Detect counterclockwise and clockwise neighbors along the bounded region. Compute distances in coordinates relative to position origin.
(3) Compute control value, next desired position defined by corresponding point $p_i(t + 1)$ along the bounded region.
(4) Move to new position $p_i(t + 1)$ along the bounded region.
(5) Get measurement of target and send it to central server.
End while

Algorithm 1: Movement tracking algorithm.

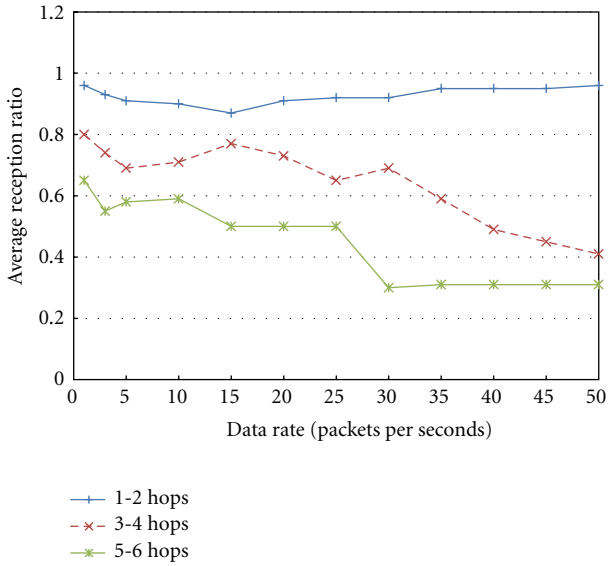


Figure 4: Patient data reception rate.

Table 1: Coverage comparison for sampling probability function.

Sample	P_{TZ} (Random)	P_{TZ} (Uniform)	P_{TZ} (Optimal)
1	0.2456	0.2564	0.2568
2	0.4327	0.4580	0.4656
3	0.5212	0.5368	0.5523
4	0.6002	0.6092	0.6257
5	0.2856	0.3059	0.4568

Table 2: Probabilistic measure of optimal placement performance.

Sample	P (> random)	P (> uniform)
1	0.4129	0.2596
2	0.5490	0.5028
3	0.7831	0.7649
4	0.8412	0.8018
5	0.9995	0.9654

varies over time because ADMR routes are dynamic. The single-hop case should be very common in clinical settings where the doctor or nurse is generally near the patient.

The numerical approach used to calculate P_{TZ} from particular wireless sensor coverage is composed of establishing initially a resolution grid of the track parameters and then counting the number of sensors occurring within each target region corresponding to a particular track position and direction. P_{TZ} is then given as the ratio of target region monitored to the total number being present in the monitoring region.

To verify the utility of this placement scheme a Monte Carlo simulation was performed. The steps for experiment included the following. For N sensors, (1) generates a random sample within monitoring region Z. (2) Generates a random sample uniformly within monitoring region Z. (3) Generates a random sample for optimal calculation of the sensor area coverage function $\gamma_z(t)$. (4) Calculates the corresponding P_{TZ} from each sampling.

The probability of performing better than uniform is then estimated as the ratio of this count and the total number of Monte Carlo simulation runs. This experiment is repeated to compare sampling from the optimal sensors that are coverage function. Table 1 shows the P_{TZ} calculated by sampling from sensor coverage using the target characteristics corresponding to each example. The values of P_{TZ}, calculated from the Monte Carlo simulation, show that, for each sample, the optimal is better than the uniform which is constrained within Z and the random case. The largest improvement was in sample 5, corresponding to the most stringent sensor detection criteria, while the least improvement was for area coverage sample 1 where uniform is close to optimal. Another table shows a probabilistic comparison of the performance of the sampled optimal sensor coverage to that of the uniform and random cases. This is shown in Table 2 which contains the numbers that represent the probability that a random sample of 50 from the optimal sensor coverage area results in a higher P_{TZ} than that of the random and uniform cases.

These observations from the numerical procedure described in this research showed two computational pieces, a genetic algorithm and a semidefinite programming algorithm approach. In actual experiment, for the samples in

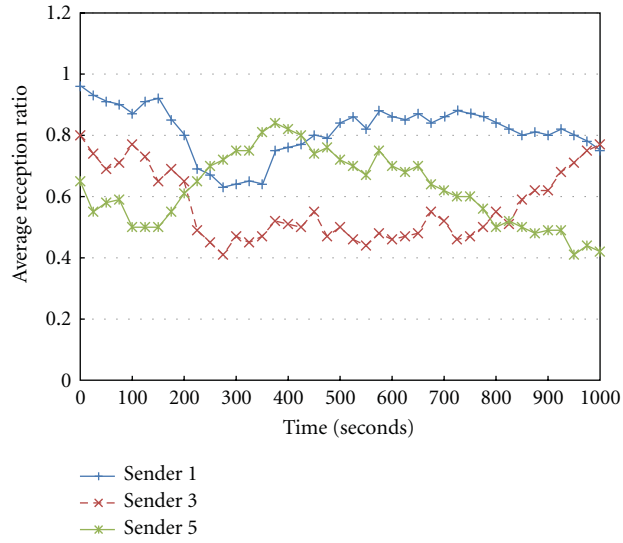


Figure 5: Effect on mobility.

this paper, the genetic algorithm consumed the majority of the computational time, 60% of the time. Following it was the semidefinite programming which consumed 25% of the time. Lastly, the placement procedure took approximately 15% of the total time usage. MATLAB software was used for the optimization procedure. The sensor positions were used as basis for the calculation of the optimization procedure for both the genetic algorithm and semidefinite programming. It is expected that computational time of the two-level optimization is relatively independent of the scale of the problem.

6.1. Effect on Mobility. Mobility of the senders or receivers of information has impact on communication reliability. As senders or receivers move in a hospital, radio link quality will vary and ADMR will create new routes. Therefore, we expect to see some data loss due to node mobility, but ideally a valid route will be maintained at all times.

In this experiment, we consider fixed nodes as patient sensors transmitting data at 5 packets per second. The senders were widely distributed throughout the building. A single receiver node attached to a laptop acted as a roaming node. The user carrying the laptop moved around the second floor of our building at a normal walking pace, pausing occasionally, entering and leaving rooms, for a duration of about 25 minutes. This movement pattern is intended to represent a doctor walking through a hospital ward.

Figure 5 shows the reception ratio for each of the 3 senders, averaged over 60-second windows. As the receiver walks around, we see the reception ratios vary over time but do not see any large dropouts or catastrophic effects due to mobility. We have also recorded the hop count and ADMR path cost for each packet and see a general correlation between improved delivery ratio and reduced path cost. These results show that ADMR deals gracefully with node movement, at least for typical mobility rates.

6.2. Low Latency Transmission. In wireless ad hoc and sensor networks, the problem of routing has received more attention than any other design and operation problem. Many wireless routing algorithms have been proposed in the last couple of decades. Flooding and broadcast routing is often necessary during the operation of the wireless network, such as to discover node failure and broadcast some information. Multicast routing, on the other hand, is very common in wireless networks, and it is used to communicate in a one-to-group fashion. Moreover, it involves wireless multicast advantage (WMA) [22] which means that, if a node transmits a packet by spending high power, it is possible that more than one node receive its transmission. Finally, unicast is always in an end-to-end fashion and it is the most common kind of routing in networks. The case of unicast routing, although a special case of multicasting, involves no wireless advantage; however, choosing a good path from source to destination requires knowledge of node and link states. This is especially the case when battery lifetime maximization is an objective. Given a selected route, nodes on this route between the source and destination who act as routers deplete their energies with each packet they forward.

6.3. Reliability. The best approach to implementing reliability is not immediately clear. Using link-by-link acknowledgment and retransmission with multicast requires additional MAC support and may incur high overhead. End-to-end reliability is highly sensitive to overall path conditions.

One approach that is worth considering makes use of redundant transmissions and coding techniques that allow data to be reconstructed on the receiver despite packet loss. We are still investigating this idea, but, to capture a rough estimate of how it would perform, we have conducted experiments where each message is simply transmitted multiple times by the sender. In this way, a receiver can recover the original data if any one of k transmitted packets is received. This approach consumes considerably more bandwidth but should yield an estimate of the improvement obtainable via more sophisticated techniques.

7. Conclusion

This paper has presented the deployment of distributed wireless network of sensors for monitoring target patients. An optimization scheme was implemented for optimal placement of sensors and movement coordination techniques within a search region given the underlying characteristics of sensors and expected targets. A movement tracking algorithm was also proposed to serve as a guide for the wireless sensor networks for optimal deployment and provide distributed detection criteria. The problem for placement of sensors was addressed as a sampling from the optimal sensor density, and a deterministic conditional sampling approach for placing individual sensors was developed and compared to random sampling. With the practical advantages of deploying sensor networks using density-based approach, it would be of clear interest to modify our model by including the upper bounds of the movement and detection range of

the wireless sensors. Broader future research includes the consideration of more complex and heterogeneous collection of sensors and the dynamic assignment of wireless sensors to different patient targets.

Despite the fact that there is keen interest in M2M technology and great value in building an efficient M2M network, M2M is still relatively new and the technology faces several significant challenges. Major challenges today, in addition to energy efficiency, are in the areas of security, privacy, reliability, robustness, latency, cost-effectiveness, software development, and standardization. Although many routing techniques look promising in terms of energy efficiency, most of these algorithms were designed for a network where nodes are stationary. While it is true that most of today's M2M applications have few mobile nodes in a network, in the near future, there will be many M2M networks consisting of hundreds of mobile nodes.

For future study, we aim to reduce or eliminate the signaling overhead of exchanging status information by some feature extraction and local estimation functions.

Acknowledgment

This research was supported by the ICT Standardization program of KCC (Korea Communications Commission).

References

[1] G. Lawton, "Machine-to-machine technology gears up for growth," *Computer*, vol. 37, no. 9, pp. 12–15, 2004.

[2] R. Kwok, "Use of (1) sensors and (2) radio frequency ID (RFID) for the national children's study," Technical Report, RTI International, 2004.

[3] IEEE 802.15.4, "Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (LR-WPANs)," IEEE, 2003.

[4] 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.

[5] S. Tilak, N.B. Abu-ghazaleh, and W. Heinzelman, "A taxonomy of wireless micro-sensor network models," in *ACM Mobile Computing and Communications Review*, vol. 6, pp. 28–36, 2002.

[6] O. Boric-Lubecke and V. M. Lubecke, "Wireless house calls: using communications technology for health care and monitoring," *IEEE Microwave Magazine*, vol. 3, no. 3, pp. 43–48, 2002.

[7] P. Várady, Z. Benyó, and B. Benyó, "An open architecture patient monitoring system using standard technologies," *IEEE Transactions on Information Technology in Biomedicine*, vol. 6, no. 1, pp. 95–98, 2002.

[8] K. Hung and Y. T. Zhang, "Implementation of a WAP-based telemedicine system for patient monitoring," *IEEE Transactions on Information Technology in Biomedicine*, vol. 7, no. 2, pp. 101–107, 2003.

[9] K. Akkaya and M. Younis, "A survey on routing protocols for wireless sensor networks," *Ad Hoc Networks*, vol. 3, no. 3, pp. 325–349, 2005.

[10] J. N. Al-Karaki and A. E. Kamal, "Routing techniques in wireless sensor networks: a survey," *IEEE Wireless Communications*, vol. 11, no. 6, pp. 6–27, 2004.

[11] R. Kwok, "Use of sensors and radio frequency ID (RFID) for the national children's study," Technical Report, RTI International, 2004.

[12] S. Ferrari, "Track coverage in sensor networks," in *Proceedings of the American Control Conference*, pp. 2053–2059, Minneapolis, Minn, USA, June 2006.

[13] S. S. Ram, D. Manjunath, S. K. Iyer, and D. Yogeshwaran, "On the path coverage properties of random sensor networks," *IEEE Transactions on Mobile Computing*, vol. 6, no. 5, pp. 446–458, 2007.

[14] A. Howard, M. Mataric, and G. Sukhatme, "Mobile sensor network deployment using potential fields: a distributed, scalable solution to the area coverage problem," in *Proceedings of the 6th International Symposium on Distributed Autonomous Robotics Systems (DARS '02)*, Fukuoka, Japan, June 2002.

[15] S. Dhillon and K. Chakrabarty, "Sensor placement for effective coverage and surveillance in distributed sensor networks," in *Proceedings of the IEEE Wireless Communications and Networking (WCNC '03)*, vol. 3, pp. 1609–1614, New Orleans, La, USA, March 2003.

[16] T. A. Wettergren and R. Costa, "Optimal placement of distributed sensors against moving targets," *ACM Transactions on Sensor Networks*, vol. 5, no. 3, pp. 1–25, 2009.

[17] M. Huff, "M2M device networking: components & strategies," Technical Report, MSI Tec, September 2007.

[18] K. Bhaskar, B. W. Stephen, and B. Ramon, "Phase transition phenomena in wireless Ad-Hoc networks," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '01)*, vol. 5, pp. 2921–2925, San Antonio, Tex, USA, November 2001.

[19] D. Ucinski, "Optimal Measurement Methods for Distributed-Parameter System Identification," CRC Press, 2005.

[20] J. G. Jetcheva and D. B. Johnson, "Adaptive demand-driven multicast routing in multi-hop wireless ad hoc networks," in *Proceedings of the ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc '01)*, pp. 33–44, Long Beach, Calif, USA, October 2001.

[21] A. Schwartz, E. Polak, and Y. Chen, RIOTS, a matlab toolbox for solving optimal control problems, 1997, http://mechatronics.ece.usu.edu/riots/.

[22] J. E. Wieselthier, G. D. Nguyen, and A. Ephremides, "Energy-efficient broadcast and multicast trees in wireless networks," *Mobile Networks and Applications*, vol. 7, no. 6, pp. 481–492, 2002.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 171738, 12 pages
doi:10.1155/2012/171738

Research Article

Logic Macroprogramming for Wireless Sensor Networks

Supasate Choochaisri, Nuttanart Pornprasitsakul, and Chalermek Intanagonwiwat

Department of Computer Engineering, Chulalongkorn University, Bangkok 10330, Thailand

Correspondence should be addressed to Chalermek Intanagonwiwat, intanago@yahoo.com

Received 3 October 2011; Accepted 17 December 2011

Academic Editor: Tai Hoon Kim

It is notoriously difficult and tedious to program wireless sensor networks (WSNs). To simplify WSN programming, we propose Sense2P, a logic macroprogramming system for abstracting, programming, and using WSNs as globally deductive databases. Unlike macroprograms in previous works, our logic macroprograms can be described declaratively and imperatively. In Sense2P, logic macroprogrammers can easily express a recursive program or query that is unsupported in existing database abstractions for WSNs. We have evaluated Sense2P analytically and experimentally. Our evaluation result indicates that Sense2P successfully realizes the logic macroprogramming concept while consuming minimal energy as well as maintaining completeness and soundness of the answers.

1. Introduction

Wireless sensor networks (WSNs) have been widely used for collecting data from environments [1–4]. However, sensor nodes are resource constrained and distributed all over the monitored area. Programming WSNs to acquire such data is notoriously difficult and tedious. Traditional WSN programming requires system programming in low-level details (e.g., wiring nesC [5] components, coordinating the program flow among nodes in a distributed manner, routing, discovering resources, accessing, and managing remote data) while maintaining low energy consumption and memory usage [6].

Several programming abstractions have been proposed to simplify WSN programming with high-level languages and to hide the low-level details from programmers [6–12]. The WSN programming abstractions have been divided into two classes: local-behavior class and global-behavior class (also called *macroprogramming* class). The abstraction in the former class simplifies the programming task of specifying the local behavior of each node for distributed computation. Local-behavior abstractions include abstract regions [11, 12] and DSN [6]. These local-behavior abstractions can efficiently hide some of the above low-level programming details but the programmers still need to write a distributed code for routing, coordinating the program flow among nodes, accessing, and managing remote data.

Conversely, the abstraction of the macroprogramming class enables expressing the global behavior of the distributed computation by programming the WSN in the large [7]. These macroprogramming abstractions can hide even more low-level programming details than the local-behavior abstractions do. In a sense, macroprogrammers take a centralized view of programming a distributed system rather than a distributed view. The macrocompiler is responsible for translating the macroprogram into a distributed version for execution.

There are two subclasses of macroprogramming abstractions: node dependent and node independent. In the node dependent subclass, a WSN is abstracted as a collection of nodes that can be simultaneously tasked within a single program. Examples of the node-dependent subclass include Kairos [7], Regiment [10], Split-C [13], SP [14], and DRN [8].

By contrast, in the node-independent subclass, a WSN is abstracted and programmed as a whole or a unit instead of several interacting nodes. Low-level programming details are completely abstracted out in this subclass as there are no longer networks or nodes in the programmer's view. Examples of this subclass include TinyDB [9] and Cougar

[15]. Both have abstracted WSNs as relational databases that are programmed or queried in a SQL-like language. This abstraction is reasonable because WSNs have also been queried for data in relation [16]. Given this database abstraction, WSN programming is reduced to database querying.

However, SQL is a pure declarative programming language for specifying what the programmer wants, not how to algorithmically obtain the desired result. Despite its simplicity, declarative programming may not be applicable to several WSN applications, especially complex tasks or queries. Undoubtedly, imperative programming (or procedural programming) is more appropriate for such complex tasks where efficient algorithmic details are application specific, unobvious, or difficult to generate automatically. Declarative and imperative programming approaches function well within their domain and complement one another. Integration of both approaches can form a powerful programming paradigm suitable for both domains.

Widely considered such integration, logic programming is the use of logic as both a declarative and imperative representation language [17]. A logic program consists of declarative sentences in the form of implications. Based on a backwards reasoning theorem prover, logic programming treats the implications as goal-reduction procedures. Logic programmers can exploit the problem-solving behavior of the theorem prover to achieve efficiency. This is similar to how imperative programmers use programs to control the behavior of a program executor. However, unlike pure imperative programs, the correctness of logic programs can be ensured with their declarative and logical interpretation.

In this paper, we propose Sense2P, a logic macroprogramming system for abstracting and programming WSNs as globally deductive databases. Unlike macroprograms in previous works, our logic macroprograms can be described declaratively and imperatively. As a result, Sense2P is highly expressive and efficient, compared to SQL-based systems.

Another advantage of logic macroprogramming is its capability to easily express a recursive program. Even though one can express a recursive query in SQL, the recursive SQL query is rather verbose (see appendix A) and unsupported in existing systems for WSNs.

Our evaluation result indicates that Sense2P can realize the logic macroprogramming concept while consuming minimal energy and maintaining completeness and soundness of the answers.

The remainder of the paper is described as follows. Section 2 reviews related work about macroprogramming and logic programming in WSNs. Section 3 describes the logic macroprogramming approach to WSNs. Then, we explain Sense2P in Section 4. Sections 5 and 6 cover our programming model and system architecture, respectively. We mathematically analyse the communication cost of our approach in Section 7 and experimentally evaluate the performance of our system in Section 8. Finally, Section 9 concludes the paper.

2. Related Work

Various macroprogramming abstractions have been proposed for several years. However, no abstraction fits all domains. We discuss the differences of our abstraction from those existing ones in this section.

Of a particular interest are Kairos [7], Regiment [10], and DRN [8]. Kairos presents the programming model that computes a set of sensor devices in parallel and provides a facility to sequentially access remote variables. Unlike Kairos, Regiment is the *spatiotemporal* macroprogramming system that is based on the concept of functional reactive programming. However, Regiment is designed for long-running queries (not well-suited for short-lived queries).

DRN is a hybrid approach between imperative programming and declarative programming. Resources and nodes are declaratively named whereas the core algorithm is imperatively programmed. Similar to DRN, Sense2P is also a hybrid approach, given that logic programming is an integration of imperative programming and declarative programming. Kairos, Regiment, and DRN are node dependent but Sense2P is node independent.

Semantic Stream [18] is a macroprogramming framework with logic programming features that allows users to pose declarative queries over semantic interpretations of sensor data. However, Semantic Stream focuses on finding available services and providing the quality of services instead of problem solving. Furthermore, it is not designed specifically for wireless sensor nodes with limited resources.

Chu et al. [6] have further developed the concept of logic programming into Snlog for programming WSNs and enabling recursive queries. Snlog, however, is designed for low-level programmers, not for application-level programmers. Unlike Sense2P programmers, Snlog programmers must write rules by focusing on local behaviors of each sensor node (instead of the global behavior as a whole). Therefore, Snlog does not support a join between different nodes. Additionally, the Snlog programmers must deal with networking details and protocols, such as routing, query disseminating, and data collecting. In summary, Sense2P is a macroprogramming approach but Snlog is not.

TinyDB [9] and Cougar [15] are probably the most cited node-independent abstractions for macroprogramming WSNs. Those approaches abstract a WSN as a relational database. Consequently, WSN programming is reduced to database querying. However, there are several limitations in the mentioned approaches.

First, supported queries in the previous works are quite limited. For example, there is only one table accessible at a time. This may not work in networks of heterogeneous sensors. In other words, their queries do not support a join between different sensor nodes. In addition, only conjunctive comparison predicates are supported, and arithmetic expressions are limited to operations of an attribute and a constant. As a result, tuple selection is inflexible. Furthermore, subqueries and column aliases are not allowed either.

Second, each sensed data item is kept as a tuple associated with each node. Constraints in the query are applied only to attributes in the same tuple as well as the same node

Table 1: Characteristics comparison.

Approach	Characteristic				
	Programming model	Abstraction level	Node dependency	Communication transparency	Recursive query
Kairos	Imperative (procedural programming)	Network level (global)	Node dependent	Yes	No
Regiment	Declarative (functional programming)	Network level (global)	Node dependent	Yes	Yes
DRN	Declarative and imperative (procedural programming with resource variable)	Network level (global)	Node dependent	Yes	No
Cougar	Declarative (SQL)	Network level (global)	Node independent	Yes	No
TinyDB	Declarative (SQL)	Network level (global)	Node independent	Yes	No
Semantic Stream	Declarative (logic programming)	Network level (global)	Node independent	Yes	No
Snlog	Declarative and imperative (logic programming)	Node level (local)	Node dependent	No	Yes
Sense2P	Declarative and imperative (logic programming)	Network level (global)	Node independent	Yes	Yes

[9]. Therefore, the constraints are local, not global. It is not designed for deriving data that is related with other data from different nodes. As a result, they cannot support a join operation.

Third, they do not support recursive queries. It is well documented that the recursive queries can improve the capability of a database [19, 20].

Finally, previous systems with a relational-database abstraction do not support a logic-based query frequently used in deductive databases and expert systems.

Unlike TinyDB and Cougar, our approach abstracts a WSN as a globally deductive database that can be logically programmed. As a result, our approach does not suffer from the above limitations.

We summarize the characteristic differences of related works in Table 1.

3. Logic Macroprogramming

Logic programming is a logic-based declarative approach to knowledge representation that allows recursive programming. Logic programming is widely used in many artificial-intelligence applications such as knowledge-based systems, expert systems, and smart information-management systems, and so forth. Prolog [17] is a de facto language for logic programming in traditional systems. Logic programming can be combined with relational databases in order to construct "deductive database" systems that support a powerful formalism and operate quickly even with very large data sets. Their powerful features include a capability to process recursive queries and their superior expressiveness over relational databases. For example, a Prolog system can be loosely coupled with a relational database system [21] to become a deductive database system (i.e., a relational database system with an inference engine). Our early work in abstracting WSNs as deductive databases has been presented in [22].

Given this deductive-database abstraction, tasks can be logically macroprogrammed in a Prolog-like language. In WSNs, each node senses the environmental data periodically or reactively. The sensed data is locally stored and viewed as a fact in our system (see Section 4 for more details). These facts are available to logic macroprogrammers as if the facts are on the centralized database. Logic macroprogrammers simply focus on what data they need (declaratively) and how to process the data (imperatively) but not on how to retrieve those data. The macroprogrammers can create facts and rules as well as inject queries into our network-transparent system as if the network is a deductive database.

Furthermore, the macroprogrammers can also write rules for deducing new facts from existing facts and rules recursively.

4. Sense2P

Sense2P is our prototype for logic macroprogramming WSNs in a Prolog-like language. Our system allows programmers to write recursive and nonrecursive rules (programs) without being concerned with low-level programming details. Additionally, Sense2P is sufficiently simple for application-level users who only want to query the system for interested data. Our programming model and system architecture are described as follows.

5. Programming Model

Briefly, our programming language in Sense2P is Prolog like. The language consists of predicates, facts, rules, and queries.

5.1. Predicate. Predicates are relations of data (or tables in the relational-database terminology). For example, a predicate temperature (NodeID, Temperature Value) describes a relation between a node identification number and a temperature value. From this example, temperature is a predicate name while NodeID and Temperature Value are

variable arguments. In general, an argument is a variable if it begins with the capital letter. Conversely, an argument is a constant if it begins with the small letter or it is a number. Predicates in Sense2P are divided into 3 categories: user defined, built-in, and sensor specific. The first predicate type includes arbitrary predicates defined by programmers. The second type includes general predicates that are already built in the system. Examples of built-in predicates are *sum* (for computing the summation of two values) and *abs* (for computing the absolute value). The last type includes only built-in predicates that are specific to sensors. The previously mentioned temperature predicate is such a sensor-specific predicate. In Sense2P, the sensor-specific predicates are processed differently from other types. We describe our processing methodology in the next subsection.

5.2. Fact. A fact is a predicate whose arguments are all constant. One may consider facts as already-existing data in the system. Facts can be instantiated in three forms: user defined, sensor generated, and rule deduced. Users can define known facts in the program such as *location* (1, 33, 45). This fact indicates that a node ID 1 is located at coordinate (33, 45). Some facts are data-sensed from sensors. For example, temperature (3, 25) is a fact that indicates a temperature of 25°C sensed by a node ID 3. Finally, facts can also be deduced from existing facts and rules.

5.3. Rule. Rules are clauses that deduce new facts from existing facts. Rules are represented as Horn clauses that contain head and body parts. An example of a rule is shown in Listing 1.

Specifically, an area has a hot spot if an arbitrary node in that area senses temperature with a value over 50 degrees. The left-hand side of a clause is called head and the right-hand side is called body. In Listing 1, the head is *hotSpotArea(AreaID)* and the body is *temperature(NodeID, Temp), Temp* > 50, *area(NodeID, AreaID)*. A rule will be satisfied only if every predicates in the body are satisfied or matched by at least one fact.

5.4. Query. Queries are questions that a user asks to retrieve data from the system. A query is represented by *?*-followed by a predicate. For example, *?-hasHotSpotArea(X)* is a query to retrieve IDs of all areas that has a hot sensor node.

Queries can be classified into 4 groups. The first group is a fact-checking query that is intended for checking the existences of certain facts. This query type is expressed by a predicate with no variable argument, for example, *?-temperature(2, 5)*. The second group includes fact-retrieving queries that are designed for retrieving all data that satisfy the fact types and constraints in the queries. This query type contains at least one variable in a predicate, for example, *?-temperature(X, Y)*. The third group consists of queries for checking whether existing rules can deduce a certain fact. It is almost like the fact-checking query except that the predicate matches with a rule (instead of a fact), for example, *?-hotSpotArea(5)*. The final group is composed of deductive queries for retrieving all data that satisfy the rules

```
hotSpotArea(AreaID):- temperature(NodeID, T)
                    , T > 50
                    , area(NodeID, AreaID)
```

Listing 1: Example of a rule.

```
danger(AreaID):- temperature(NodeID,T)
               , T > 80
               , area(NodeID, AreaID).
danger(AreaID):- humidity(NodeID, H)
               , H < 40
               , area(NodeID, AreaID)
               , adjacent(AreaID, AdjAreaID)
               , winddir(AdjAreaID, AreaID)
               , danger(AdjAreaID).
Query: ?-danger(X).
```

Listing 2: Example of recursive query rules.

in the queries, for example, *?-hotSpotArea(X)*. A query will be recursive if its predicate matches with a recursive rule whose body contains the same predicate name as that in its head. For example, one can write recursive rules for detecting sensor nodes in danger as shown in Listing 2.

The first rule is a base case of a recursive program that describes properties of areas in danger. An area will be in danger if there is at least one node (in that area) whose sensed temperature is greater than 80 degrees (which may be on fire). The second rule is a recursive case. Basically, an area will be in danger if there is at least one node (in the area) whose sensed relative humidity is lower than 40 % (the weather is dry) and there is wind from the adjacent area in danger. When the Sens2P program starts, a user can define facts and rules before injecting a query into the system.

6. System Architecture

Sense2P consists of two major components: the query processing engine and the data-gathering engine (Figure 1). The query processing engine resides on the base station while the data-gathering engine resides on each wireless sensor node.

6.1. Query Processing Engine. The query processing engine is crucial for logic macroprogramming WSNs. The main tasks are to interpret a user program (consisting of facts, rules, and queries) and to process queries to find satisfying answers.

Sense2P query processing engine consists of three main components: a compiler unit, a run-time processing unit, and a network interface unit. The compiler unit parses a logic macroprogram into a compiled code that runs on the run-time processing unit. The run-time processing unit is required to treat sensor-specific facts differently from those of other ordinary facts. Sensor-specific facts are data

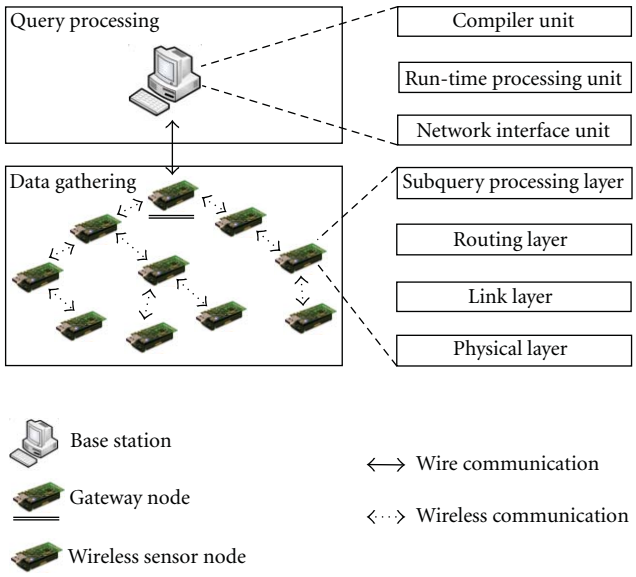


Figure 1: System Architecture.

locally sensed and stored by sensor nodes in the network. Processing queries related to these facts requires special attention because unnecessary data transmissions are costly in WSNs.

In this paper, we consider three previously proposed schemes for query processing in deductive databases. These schemes include top-down, bottom-up, and Prolog-style evaluation approaches [21].

Prolog-style systems (coupled with database systems) are similar to the top-down systems in a sense that their execution starts from the goal and a query can be solved by executing each subgoal until the deduced facts match the goal. However, Prolog-style systems produce answers one tuple at a time whereas top-down methods produce one set at a time without in-order execution of subgoals.

Conversely, the bottom-up methods start from existing facts and attempt to deduce new facts from rules that are related to the query. Only facts that match with the goal of the query are selected as the answers. We refer to [21] for more information of each implementation scheme.

Many works suggest that the bottom-up methods have many advantages over top-down methods in traditional deductive database systems [20, 21, 23]. However, in wireless sensor networks, we argue that top-down and Prolog-style approaches are more appropriate.

Understandably, facts in a wireless sensor network are data that are sensed from an environment. They are locally kept within sensing nodes and sent to the base station only when requested. To process a query in a bottom-up manner, we need facts from all relevant nodes so that new facts can be globally deduced. Therefore, each node may be required to send its data to a rendezvous point (e.g., a base station, an inference engine) for such a deduction. Undoubtedly, the mentioned mechanism consumes excessive energy. To reduce this energy consumption, only relevant data should be delivered.

```
hotObject(Obj, AreaID):- detect(Obj, AreaID)
                       , temperature(Obj, T)
                       , T > 50.
```

Listing 3: Example of a rule that two predicates related to each other with Obj.

However, it is not easy for a node to selectively send relevant data without knowing priori what all other nodes have. A fact in a node may be relevant simply because another fact from another node happens to have a certain value.

Conversely, the top-down approach can use information from a query to suppress irrelevant facts from being sent. For example, a predicate *detect(ObjectID, AreaID)* in a system means a sensor node can detect an object with the identification number *ObjectID* in the region *AreaID*. When a user injects a query *?-detect(oiltank,X)*, only sensor nodes that can detect an object named *oiltank* will send answers back. Other nodes are suppressed.

Furthermore, we can even suppress unnecessary query forwarding and redundant answering. The benefit is evident in fact-checking queries, such as *?-detect(oiltank, area70)*. In our system, only the first node detecting *oiltank* in area 70 will reply, although there may be other nodes (in the same area) that detect the same event. This is reasonable, given that one's reply about the fact existence is sufficient to satisfy the query. Therefore, the first detecting node does not need to forward the query further. Thus, there is no other replier (see Section 6.2 for more details).

In addition, we can use an answer set from the previous subgoal to filter out (or suppress) the irrelevant facts of the next subgoal. For example, consider the rule in Listing 3.

When a user injects a query *?-hotObject(X, area70)*, the system will match the query with the above rule. Therefore, the variable *AreaID* in the rule will be bound with the constant *area70*. Then, the system will attempt to match each predicate in the body of the rule. Each body predicate becomes a subquery that needs to be satisfied.

In this example, the first subquery is *detect(ObjectID, area70)*. This subquery is disseminated into the network. Only eligible repliers are nodes with facts or rules that match the subquery. Others are suppressed. Consequently, only objects in *area70* will be bound to the variable *ObjectID*. Then, each *ObjectID* will be used to bind *temperature(ObjectID, Temp)* predicate and can be used to filter out or suppress irrelevant facts from being sent. Furthermore, we can also use a constraint *Temp > 50* as another filter before injecting a subquery *temperature(ObjectID, Temp)* to the network.

Due to these filtering techniques, this top-down approach can significantly reduce the consumption of energy that is limited in wireless sensor networks [16]. Therefore, the Prolog-style top-down approach is used and combined with our filtering techniques in this paper.

In our system, most relevant facts are pulled from the network except the persistent ones that do not change over

time. The persistent facts can be cached or kept in the backend database of the inference engine for future uses.

Additionally, we propose a *superset-caching* technique to reduce even more energy consumption. With this technique, Sense2P will cache the answer set of the first-timer query (that requires sensor-specific predicates) for future use. If a user injects a new query, Sense2P will check whether one of the previous queries is a superset of the current query or not. If a previous query is the superset, Sense2P will search the satisfying answers in the cache rather than in the network. Otherwise, the query will be disseminated into the network.

For example, an answer set of a query *?-light(X,Y)* is a superset of a query *?-light(3,X)* because *light(X,Y)* contains every possible answers in the network. Conversely, an answer set of a query *?-light(5,X)* or *?-light(X,20)* is not a superset of *?-light(3,X)*. However, this caching technique is not perfect, especially when the data in the environment is volatile. Therefore, a timer to flush cache may be needed. Nevertheless, the flushing period is a trade-off between the data freshness and the energy consumption.

Finally, the network interface unit is responsible for disseminating queries or subqueries into the network. The queries are transformed into a format known in sensor networks, serialized, and sent into the network. The unit is also responsible for receiving answers from the network. This requires deserialization and transformation of messages back into the Prolog-like predicates.

Our approach works well on both recursive and nonrecursive queries. The query-processing flow is illustrated in Figure 2.

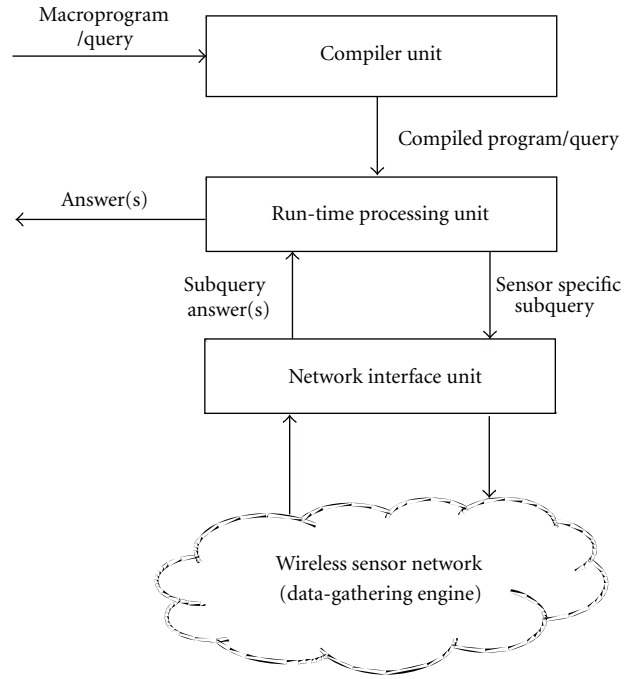


Figure 2: Query processing flow.

6.2. Data-Gathering Engine. Data-gathering engine is responsible for finding answers that are relevant to injected queries. This engine consists of the routing layer, the query-processing layer, the link layer, and the physical layer. However, our work simply focuses on the routing layer and the query processing layer. Both mentioned layers are handled by our LogicQ sub-system.

LogicQ is the underlying subsystem for subquery processing in Sense2P. Running on each sensor node, LogicQ is implemented in TinyOS [24], the operating system for the sensor mote platform. The functionality of LogicQ is to find answers for each subgoal that needs data from wireless sensor networks and minimizes the energy consumption.

When Sense2P starts up, LogicQ constructs a routing tree for disseminating subqueries from the base station to sensor nodes and for collecting answers that satisfy the subqueries. We use a drain routing tree of TinyOS as our routing tree. A root of the tree is the gateway node connected to the base station. Each node can have many child nodes but only one parent node. When disseminating the queries, we simply forward queries along the drain tree except suppressible queries (i.e., no longer necessary to be forwarded because the queries have been satisfied). Then, when answers are ready, each node sends its answers back along this routing tree.

Subgoal predicates that LogicQ is responsible to support are sensor-specific built-in predicates. Such predicates include predicates that are related to specific functions of sensor nodes' capabilities, such as *temperature(NodeID, Temp)* for sensing the temperature and *connect(NodeID1, NodeID2)* for checking connectivity, and so forth. These built-in functions are defined prior to the system installation. The Sense2P's inference engine will solve the subgoals that require sensor-specific predicates by injecting subqueries that correspond to the subgoals into the network.

In the programming model subsection, we classify queries into 4 groups. However, in this lower layer, there are only facts in sensor nodes. As a result, sensor-specific subqueries are classified into two types, one for existence checking and another for retrieving all satisfied predicates. To check an existence of a fact, every argument in this first query type is constant and the answer is only true or false (e.g., *detect(oiltank, area 70)* whereby *oiltank* and *area 70* are constant). Therefore, this type of query is not necessarily disseminated to all sensor nodes. If only one node has a fact that satisfies the query, the system does not need answers from other nodes. Conversely, the second query type requires at least one variable as an argument. For example, *detect(ObjectID, area 70)* contains a variable *ObjectID* and a constant *area70*. Hence, the system will find every possible answer of ObjectID that is detected in *area70*. It is necessary to disseminate this type of queries to all nodes.

The subquery processing algorithm for each sensor node can be written as a pseudocode in Algorithm 1. Once receiving an existence-checking query (Line 1), a sensor node checks its facts locally first whether it has a fact that satisfies the query or not (Line 2). If a sensor node has a satisfying fact, it will send an answer *true* to its parent immediately (Line 3). Given that one answer is sufficient for this query

```
if subquery is checking existence then
    if have local satisfied fact then
        send answer up to parent;
    else
        forward query to children;
    end
else if subquery is asking all Satisfied value then
    forward query to children;
    if have local satisfied fact then
        send answer up to parent;
    end
end
```

Algorithm 1: Subquery processing algorithm.

type, the replying node does not further forward the query. Otherwise, it will forward the query to its children (Line 4-5).

If the query type requires all satisfied answers (Line 6), a sensor node will forward the query immediately (Line 7). Regardless of the local existence of the satisfying facts, the system still needs satisfying answers from all sensor nodes. After the query is forwarded, the node checks for local satisfying answers. If it has one, it will send the answer up to its parent (Line 8-9).

7. Cost Analysis

In this section, we analyse the communication cost of bottom-up and top-down schemes. In all cases, we assume uniform distribution of facts in the network of M nodes. There are n subgoals in a query. $|G_i|$ represents a number of all available facts of the ith subgoal.

7.1. Cost of Bottom-Up Processing Scheme. Understandably, in the bottom-up scheme, all facts in the network must be sent to the central base station. Therefore, the total communication cost consists of rule dissemination cost, query dissemination cost, and facts retrieval cost.

The cost of rule dissemination depends on a number of subgoals because many subgoals increase the message size. For simplicity, we assume that overhead incurred by one subgoal equals to one message. Therefore, the dissemination cost of a rule with n-subgoal to M nodes in the network is

$$C_{\text{rule}} = nM. \tag{1}$$

The cost of query dissemination is obvious that a query is disseminated to all M nodes in the network. Therefore, the dissemination cost of a query is

$$C_{\text{query}} = M. \tag{2}$$

Finally, the cost of facts retrieval equals to the cost of sending all facts of each subgoal in the network to the base station. Let $|G_i|$ be a number of facts related to ith subgoal. Therefore, the facts retrieval cost is

$$C_{\text{fact}} = \overline{D}\sum_{i=1}^{n}|G_i|, \tag{3}$$

where $\overline{D}$ is an average distance from arbitrary node to the base station.

Therefore, from (1), (2), and (3), the total communication cost of the bottom-up processing scheme is

$$\begin{aligned} C_{\text{bottom-up}} &= C_{\text{rule}} + C_{\text{query}} + C_{\text{fact}} \\ &= nM + M + \overline{D}\sum_{i=1}^{n}|G_i| \\ &= (n+1)M + \overline{D}\sum_{i=1}^{n}|G_i|. \end{aligned} \tag{4}$$

7.2. Cost of Top-Down Processing Scheme. In our top-down processing scheme, we perform the join-computation process at the central base station and distributively collect only needed facts from the network.

In retrieving all satisfying answers, our scheme incurs broadcasting a subquery for each subgoal. However, only the selected facts (that satisfy the constraints caused by all previous subgoals) for that subquery are sent back to the base station. In other words, previous satisfied subgoals can suppress many unrelated facts in the network.

A rule, in this scheme, is not necessary to be disseminated into the network because the base station only disseminates a subquery of a subgoal into the network at a time. This kind of subquery is certainly a predicate that each node priori knows before deployment. Therefore, the total communication cost consists of subqueries dissemination cost and fact retrieval cost.

The cost of subqueries dissemination depends on answers of previous subgoals. These answers are used to filter irrelevant facts of the next subgoal.

Let σ_{i_1,i_2} be a selectivity factor to select facts of i_2th subgoal after solving 1st to i_1th subgoal. For example, if $\sigma_{2,3}$ equals 0.05, after solving the 1st and 2nd subgoal, only 5 percent of facts related to the 3rd subgoal are sent back to the base station. $\sigma_{-1,0}$ and $\sigma_{0,1}$ equal to 1.

The number of subqueries for the ith subgoal equals to a number of all distinct answers from the $(i-1)$th subgoal. Therefore, the number of subqueries for the ith subgoal equals to $\mu_{i-1}(\sigma_{i-2,i-1}|G_{i-1}|)$, where μ_i is a distinct factor for answers from ith subgoal and μ_0 equals to 1. For example, if there are 10 answers from several nodes but there are only 2 distinct values, the distinct factor equals to 0.2 in this case. Each subquery is disseminated to M nodes in the network. That is

$$C_{i\text{th subquery}} = M(\mu_{i-1}(\sigma_{i-2,i-1}|G_{i-1}|)). \tag{5}$$

For n subgoals, the cost of n subqueries is

$$\begin{aligned} C_{\text{subqueries}} &= \sum_{i=1}^{n} C_{i\text{th subquery}} \\ &= M\sum_{i=1}^{n}(\mu_{i-1}(\sigma_{i-2,i-1}|G_{i-1}|)). \end{aligned} \tag{6}$$

The cost of subgoal fact retrieval is similar to (3) except only a portion of facts are selected and sent back to the base station. Therefore, the subgoal fact retrieval cost is

$$C_{\text{subgoal-fact}} = \overline{D}\sum_{i=1}^{n}\sigma_{i-1,i}|G_i|. \tag{7}$$

From (6) and (7), we derive the total communication cost of the top-down processing scheme,

$$\begin{aligned} C_{\text{top-down}} &= C_{\text{subqueries}} + C_{\text{subgoal-fact}} \\ &= M\sum_{i=1}^{n}\left(\mu_{i-1}\left(\sigma_{i-2,i-1}|G_{i-1}|\right)\right) \\ &\quad + \overline{D}\sum_{i=1}^{n}\sigma_{i-1,i}|G_i|. \end{aligned} \tag{8}$$

Noticeably, our approach already includes the cost of producing all answers and the cost of sending all answers to the base station.

Note that the cost of top-down processing scheme is lower than the cost of bottom-up processing scheme when the selectivity factor is low whereas there are superfluous facts in the network. These characteristics of selectivity factor and number of facts are norm in anomaly detection applications that values exceeding defined constraints are rarely found. In the next section, we evaluate the performance of our implemented system.

8. Evaluation

To evaluate our system performance, we write and inject various queries into Sense2P that is connected to TOSSIM [25], a TinyOS simulator. For viability testing, the result is compared with that of 3 other approaches: TinyDB, bottom-up, and simplified Sense2P (a simple integration of an existing Prolog system and LogicQ without other Sense2P features). This experiment shows the impact of Sense2P features that are specifically designed for WSNs. These features include query suppressing, data filtering, and superset caching.

8.1. Performance Metrics. In this section, we use 3 metrics for performance comparison: completeness, soundness, and communication cost. Completeness is the ratio of retrieved answers to the total existing answers in the networks. This indicates whether our system can successfully retrieve all existing answers or not. Similarly, we measure the soundness by the ratio of the relevant answers retrieved to the total retrieved answers. The soundness metric is for assuring that our system does not retrieve irrelevant answers. In this evaluation, the communication cost is measured by the number of sent messages in the system. Communication cost indicates the amount of energy consumed. For a system to be viable for WSNs, the communication cost of that system must be minimized.

8.2. Simulation Environment. We implement Sense2P on TinyOS and simulate each sensor node on TOSSIM. We assume reliable communication (i.e., no packet loss because of bit errors or collisions) to discard the problem caused by radio transmissions. In our simulation, each sensor node can sense the temperature of its environment. The temperature values are randomly assigned between 20 and 80°C. We set up the simulation such that 5 percent of nodes sense the temperature value over 50°C.

8.3. Top-Down versus Bottom-Up. In this subsection, we conduct two experiments in order to compare the communication cost between the top-down evaluation of Sense2P and the traditional bottom-up evaluation. In the first experiment, we inject three temperature queries into the network of temperature sensors. The first query is for checking the existence of a predicate. All arguments in the query are constant. The second query contains one constant argument and one variable argument. In the third query, all arguments are variable. This experiment is quite simple, given that these queries are satisfied by facts, not rules.

Our message counts in answering three mentioned queries are compared with that of the bottom-up method. Regardless of the argument types, the bottom-up approach always incurs a certain amount of messages sent because all facts must be delivered to the base station (see Figure 3(a)).

Sense2P will significantly reduce the communication cost if the query contains at least one constant argument to suppress irrelevant answers. However, Sense2P will incur the communication cost similar to that of the bottom-up approach if all arguments in the query are variable. Understandably, all facts are required in order to answer such a nonconstant query under our investigated scenarios.

Nevertheless, in our simulation, we assume all nodes are equipped with the same sensor type. If the sensor nodes are heterogeneous, Sense2P will still reduce the communication cost significantly even with the nonconstant query because only relevant nodes with the matched sensing capability will send back the data, unlike the bottom-up approach that needs all facts and filters out by the inference engine at the base station.

In the second experiment, we program the rule in Listing 1 on Sense2P and inject two queries into the system: constant type and nonconstant type. Both queries must be satisfied by the mentioned rule. Expectedly, the result in Figure 3(b) indicates that our system outperforms the bottom-up approach regardless of the query types (including the nonconstant type).

However, our savings in nonconstant queries can still be improved. In our implementation, each answer from the previous subgoal is used to bind the variable in the current subgoal. The number of subqueries for the current subgoal depends on the number of answers from the previous subgoal because each answer may bind the variable with a different value.

As the number of nodes is increased, the number of answers for the previous subgoal is also increased. Consequently, the number of subqueries for the current subgoal is unavoidably increased. This causes more messages sent into larger networks. However, this problem can be solved

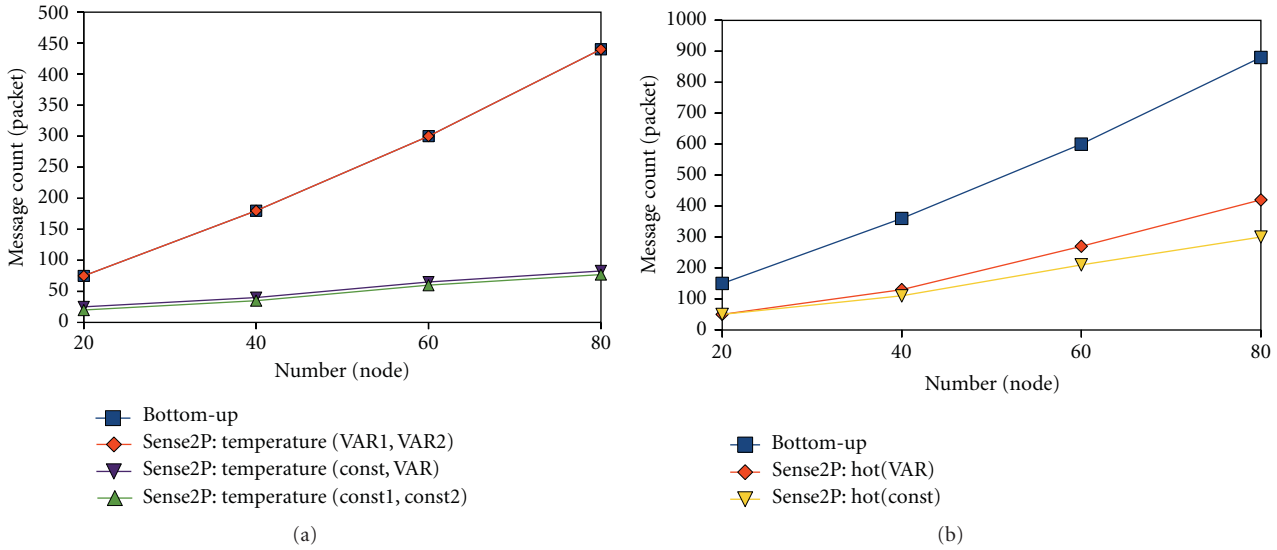


Figure 3: Comparison between traditional bottom-up evaluation and Sense2P top-down evaluation. (a) Simple predicate, temperature(NodeID, T). (b) Rule, hot(AreaID): temperature(NodeID,T), T > 50, area(NodeID, AreaID).

by sending only one subquery for the current subgoal with a list of different values that are bound with the variable. Nevertheless, we have not yet implemented this optimization in this paper. We intend to further explore this technique and other optimization approaches in our future work.

8.4. Query Suppression. In this subsection, we conduct an experiment for comparing Sense2P with TinyDB because TinyDB is also a node-independent macroprogramming paradigm. Given that TinyDB does not support many query types that Sense2P can (see Section 2), we only focus on queries that both approaches can perform.

We inject two temperature queries into the network. One is with two variable arguments for retrieving all temperature facts. Another is with two constant arguments for checking the existence of a temperature fact.

Expectedly, the message count of Sense2P is significantly smaller than that of TinyDB, especially in the constant query (Figure 4). Understandably, the efficiency of Sense2P is due to its query suppression. Sense2P suppresses (does not forward) queries that are already satisfied whereas TinyDB always sends queries to all nodes. The efficiency of Sense2P will be more evident if we measure the byte count instead of the message count, given that the message size of Sense2P is also smaller than that of TinyDB.

However, we are surprised that TinyDB sends more query messages when there are more conditions in the WHERE clause of the query. Our query with two constant arguments corresponds to a SQL query with two conditions in the WHERE clause. Evidently, TinyDB sends more query messages in the constant query than in the nonconstant query.

8.5. Data Filtering. In this subsection, we analyze the impact of data filtering on the performance of Sense2P. We program the rule in Listing 1 and inject two queries into the Sense2P system: a constant type and a nonconstant type. After that, we disable the data filtering feature in Sense2P and repeat the experiment.

Not surprisingly, Sense2P with data filtering performs better than Sense2P without data filtering in both query types (Figure 5). Data filtering is undoubtedly beneficial to the constant query. However, one may wonder how the filtering technique improves the performance of the nonconstant query. Such improvement is possible in Sense2P when the nonconstant query must be satisfied by a rule, especially the rule whose body contains a constraint to a predicate's variable argument. For example, the rule hot(AreaID), in Listing 1, contains a constraint, *Temp > 50*. In the body of the rule, the temperature predicate is the first subgoal whereas the above constraint is the second subgoal. Traditionally, a Prolog system resolves this rule from left to right. Therefore, the traditional system must retrieve every possible value of the temperature predicate before filtering the irrelevant answers with that constraint. However, in Sense2P, our runtime processing engine binds that value constraint to the temperature predicate before sending the first subquery into the network. Therefore, the number of answers for the first subgoal is reduced. Consequently, the number of messages sent in the system is also reduced.

8.6. Superset Caching. To study the impact of superset caching, we inject two set of temperature queries into Sense2P with and without superset caching. A temperature query is injected every 20 seconds in our experiment. The first 3 queries of the first set are in the form of temperature (const, VAR) where const is different for each query. The remaining queries of the first set are in the form of temperature (const1, const2). Some remaining queries are the subsets of the first 3 queries. When they are, they do not incur any radio transmission because Sense2P can search their answers in the superset caching (Figure 6(a)).

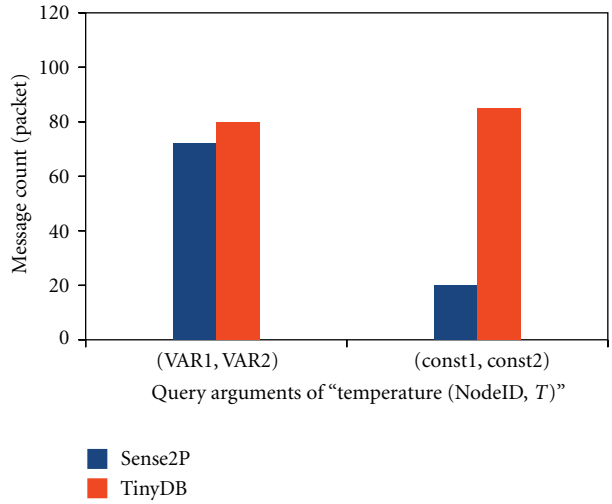


Figure 4: Comparison of Sense2P with query suppression and TinyDB with no query suppression.

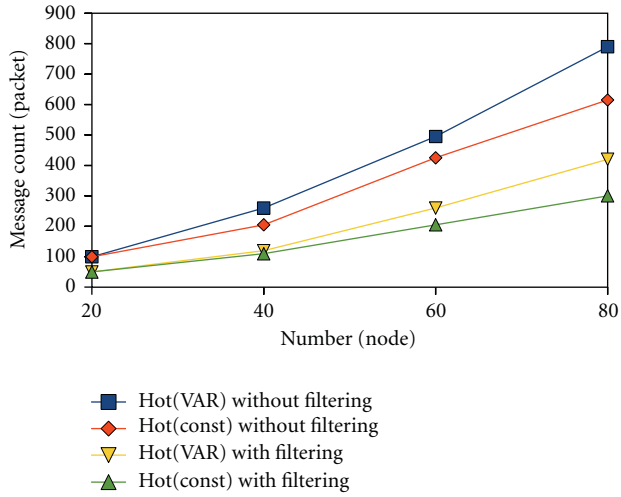


Figure 5: The impact of data filtering technique.

The impact of the superset caching is more significant when the superset is larger. This is evident in the second set of queries. The first 3 queries of the second set are in the form of temperature (VAR1, VAR2). The remaining queries contain at least one constant argument. Given that the first query is the superset of all remaining queries, there is no radio transmission necessary for answering these queries (Figure 6(b)). Consequently, the energy consumption is significantly reduced.

However, the superset caching has a trade-off issue. If the sensed data of the environment is frequently changed, the cached answer will be stale and useless. Thus, similar to most caching techniques, the superset caching is associated with an application-specific expiration timer or data popularity for flushing stale cached data. Understandably, the flushing rate is a trade-off between data freshness, storage size, and energy consumption. The cache replacement policy is out of scope of this work.

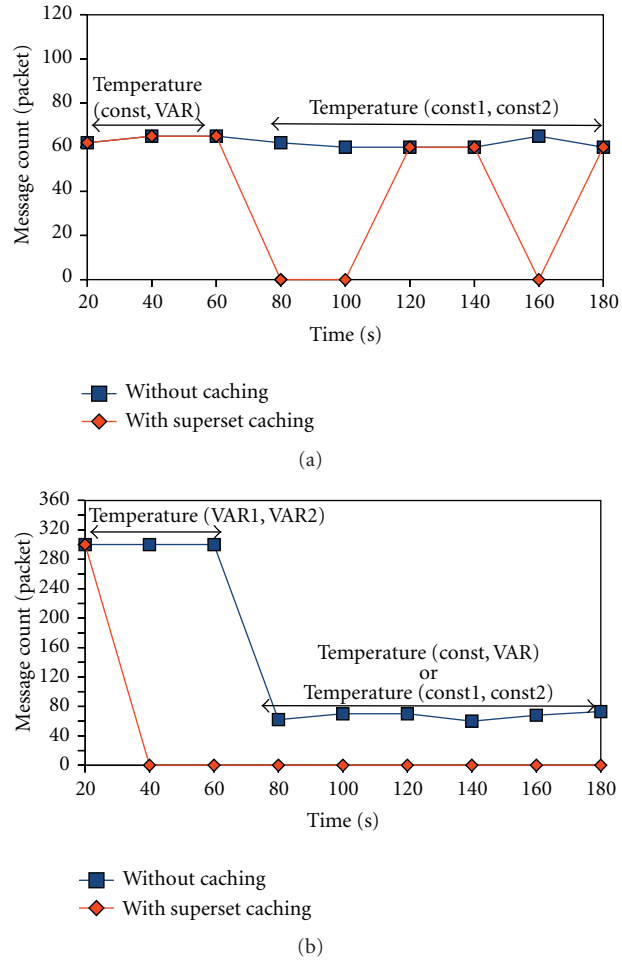


Figure 6: The impact of superset caching.

8.7. Completeness and Soundness. Due to the reliable communication in our experiment, there is no packet loss. Sense2P can achieve 100% completeness and 100% soundness under investigated scenarios because of our filtering and suppressing techniques. In practice, there would be some loss due to interference, collision, and congestion. However, to handle loss in the network is out of the scope of this work which mainly focuses on the programming language perspective.

9. Conclusion

This paper proposes a logic node-independent macroprogramming approach for abstracting, programming, and using WSNs as globally deductive databases. Unlike macroprograms in previous works, our logic macroprograms can be described declaratively and imperatively. Furthermore, logic macroprogrammers can easily express a recursive program or query that is unsupported in existing database abstractions for WSNs.

To efficiently process queries and their subqueries (either recursive or nonrecursive), the top-down approach is more appropriate than the bottom-up approach. This is due to

```
WITH RECURSIVE ancestor(anc, desc) AS (
   ( SELECT par AS anc, child AS desc FROM parent )
   UNION
   ( SELECT ancestor.anc, parent.child AS desc
      FROM ancestor, parent
      WHERE ancestor.desc = parent.par ) )
SELECT anc FROM ancestor WHERE desc="John"
```

Listing 4: SQL programming to solve Ancestors problem and Listing.

```
ancestor(anc, desc):- parent(anc, desc).
ancestor(anc, desc):- parent(anc, X), ancestor(X, desc).
?-ancestor(anc, "John").
```

Listing 5: Logic programming to solve Ancestors problem.

its capability to bind subgoal arguments that can be used to reduce communication cost by suppressing irrelevant answers and already satisfied subqueries from being sent or forwarded.

Finally, our evaluation results indicate that Sense2P can significantly reduce energy consumption while maintaining 100% completeness and soundness under our investigated scenarios.

Appendix

Recursive Query in SQL Language and Logic Programming Language

In this section, we describe the verboseness of a SQL language in expressing a recursive query. We also show the conciseness of a logic programming language in expressing the same recursive query for comparison.

To simplify the comparison, we use a well-known recursive query example: the Ancestors problem. Given a set of parent(par, child) relations, we would like to find all ancestors of "John".

The SQL program in Listing 4 demonstrates how to solve this problem.

Concisely, we can solve the same problem in logic programming as in Listing 5.

The logic program is so much shorter and easier to express. The advantage will be even more if the problem is more complex.

Acknowledgments

The authors would like to thank Rawin Youngnoi and UbiNET research group members for their valuable suggestions and effort to improve this work. This work was supported by the Thailand Research Fund (TRF) under Grant MRG5080449 and the CU CP Academic Excellence Scholarship from Department of Computer Engineering, Faculty of Engineering, Chulalongkorn University.

References

[1] C. Gui and P. Mohapatra, "Power conservation and quality of surveillance in target tracking sensor networks," in *Proceedings of the 10th Annual International Conference on Mobile Computing and Networking (MobiCom '04)*, pp. 129–143, ACM, New York, NY, USA, 2004.

[2] T. He, S. Krishnamurthy, J. A. Stankovic et al., "Energy-efficient surveillance system using wireless sensor networks," in *Proceedings of the 2nd International Conference on Mobile Systems, Applications, and Services (MobiSys '04)*, pp. 270–283, ACM, New York, NY, USA, 2004.

[3] A. Mainwaring, D. Culler, J. Polastre, R. Szewczyk, and J. Anderson, "Wireless sensor networks for habitat monitoring," in *Proceedings of the 1st ACM International Workshop on Wireless Sensor Networks and Applications (WSNA '02)*, pp. 88–97, ACM, New York, NY, USA, 2002.

[4] N. Xu, S. Rangwala, K. K. Chintalapudi et al., "A wireless sensor network for structural monitoring," in *Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems (SenSys '04)*, pp. 13–24, ACM, New York, NY, USA, November 2004.

[5] D. Gay, P. Levis, E. Brewer, R. Von Behren, M. Welsh, and D. Culler, "The nesC language: a holistic approach to networked embedded systems," in *Proceedings of the ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI '03)*, pp. 1–11, ACM, New York, NY, USA, June 2003.

[6] D. Chu, L. Popa, A. Tavakoli et al., "The design and implementation of a declarative sensor network system," in *Proceedings of the 5th International Conference on Embedded Networked Sensor Systems (SenSys '07)*, pp. 175–188, ACM, New York, NY, USA, 2007.

[7] R. Gummadi, N. Kothari, R. Govindan, and T. Millstein, "Kairos: a macro-programming system for wireless sensor networks," in *Proceedings of the 20th ACM Symposium on Operating Systems Principles (SOSP '05)*, pp. 1–2, ACM, New York, NY, USA, 2005.

[8] C. Intanagonwiwat, R. K. Gupta, and A. Vahdat, "Declarative resource naming for macroprogramming wireless networks of embedded systems," in *ALGOSENSORS, Lecture Notes in Computer Science*, vol. 4240, pp. 192–199, Springer, 2006.

[9] S. R. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, "TinyDB: an acquisitional query processing system for sensor networks," *ACM Transactions on Database Systems*, vol. 30, no. 1, pp. 122–173, 2005.

[10] R. Newton, G. Morrisett, and M. Welsh, "The regiment macroprogramming system," in *Proceedings of the 6th International Conference on Information Processing in Sensor Networks (IPSN '07)*, pp. 489–498, ACM, New York, NY, USA, 2007.

[11] M. Welsh and G. Mainland, "Programming sensor networks using abstract regions," in *Proceedings of the 1st Conference on Symposium on Networked Systems Design and Implementation (NSDI '04)*, p. 3, USENIX Association, Berkeley, Calif, USA, 2004.

[12] K. Whitehouse, C. Sharp, E. Brewer, and D. Culler, "Hood: a neighborhood abstraction for sensor networks," in *Proceedings*

of the 2nd International Conference on Mobile Systems, Applications, and Services (MobiSys '04), pp. 99–110, ACM, New York, NY, USA, 2004.

[13] A. Krishnamurthy, D. E. Culler, A. Dusseau et al., "Parallel programming in split-C," in *Proceedings of the ACM/IEEE Conference on Supercomputing (Supercomputing '93)*, pp. 262–273, ACM, New York, NY, USA, November 1993.

[14] C. Borcea, C. Intanagonwiwat, P. Kang, U. Kremer, and L. Iftode, "Spatial programming using smart messages: design and implementation," in *Proceedings of the 24th International Conference on Distributed Computing Systems (ICDCS '04)*, pp. 690–399, IEEE Computer Society, Washington, DC, USA, 2004.

[15] Y. Yao and J. Gehrke, "The cougar approach to in-network query processing in sensor networks," *SIGMOD Record*, vol. 31, no. 3, pp. 9–18, 2002.

[16] C. Intanagonwiwat, R. Govindan, and D. Estrin, "Directed diffusion: a scalable and robust communication paradigm for sensor networks," in *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking (MobiCom '00)*, pp. 56–67, ACM, New York, NY, USA, 2000.

[17] I. Bratko, *Prolog Programming for Artificial Intelligence*, Addison-Wesley Longman Publishing, Boston, Mass, USA, 1986.

[18] K. Whitehouse, F. Zhao, and J. Liu, "Semantic Streams: a framework for composable semantic interpretation of sensor data," in *Wireless Sensor Networks*, vol. 3868 of *Lecture Notes in Computer Science*, pp. 5–20, Springer, 2006.

[19] F. Bancilhon and R. Ramakrishnan, "An amateur's introduction to recursive query processing strategies," *SIGMOD Record*, vol. 15, no. 2, pp. 16–52, 1986.

[20] Y. K. Hinz, "Datalog bottom-up is the trend in the deductive database evaluation strategy," Tech. Rep. INSS 690, University of Maryland, 2002.

[21] K. Ramamohanarao and J. Harland, "An introduction to deductive database languages and systems," *The VLDB Journal*, vol. 3, no. 2, pp. 107–122, 1994.

[22] S. Choochaisri and C. Intanagonwiwat, "A system for using wireless sensor networks as globally deductive databases," in *Proceedings of the IEEE International Conference on Wireless & Mobile Computing, Networking & Communication (WIMOB '08)*, pp. 649–654, IEEE Computer Society, Washington, DC, USA, 2008.

[23] R. Ramakrishnan and S. Sudarshan, "Top-down vs. bottom-up revisited," in *Proceedings of the International Logic Programming Symposium*, pp. 321–336, MIT Press, 1991.

[24] P. Levis, S. Madden, J. Polastre et al., "Tinyos: an operating system for sensor networks," in *Ambient Intelligence*, Springer, 2004.

[25] P. Levis, N. Lee, M. Welsh, and D. Culler, "Tossim: accurate and scalable simulation of entire tinyos applications," in *Proceedings of the 1st International Conference on Embedded Networked Sensor Systems (SenSys '03)*, pp. 126–137, ACM, New York, NY, USA, 2003.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 460754, 11 pages
doi:10.1155/2012/460754

Research Article

Energy-Efficient Fire Monitoring over Cluster-Based Wireless Sensor Networks

Young-guk Ha,[1] Heemin Kim,[1] and Yung-cheol Byun[2]

[1] *Department of Computer Science and Engineering, Konkuk University, Seoul 143-701, Republic of Korea*
[2] *Department of Computer Engineering, Jeju National University, Jeju-si, Jeju 690-756, Republic of Korea*

Correspondence should be addressed to Yung-cheol Byun, ycb@jejunu.ac.kr

Received 12 November 2011; Accepted 11 January 2012

Academic Editor: Tai Hoon Kim

Uncontrolled fires occurring in wild areas cause significant damage to natural and human resources. Many countries are looking for ways to fight forest fires at an early stage using sensor networks, by integrating IT technologies. Studies in the fire-related sensor network field are broadly classified into efficient processing of fire data on sensor nodes and energy efficiency during communications among wireless sensor nodes in case of fire. Most studies of sensor network energy efficiency so far mainly focus on extending the connectivity of the entire network and minimizing isolated nodes by applying power evenly to each sensor node through efficient cluster-based routing. This paper proposes an energy-efficient fire monitoring protocol over cluster-based sensor networks. The proposed protocol dynamically creates and reorganizes the sensor network cluster hierarchy according to the direction of fire propagation over the sensor network clusters. This paper also presents experimental results to show that the proposed protocol is more energy efficient than fire monitoring with existing cluster-based sensor network protocols.

1. Introduction

The world is facing many risks caused by natural disasters such as forest fires, floods, and abnormal climate changes. IT scientists are studying ways of effectively solving such problems by looking at risks at an earlier stage. This study was conducted on a new network for detecting risk factors and rapidly responding to these problems. Wireless sensor networks are regarded as the best systems for applications in those environments. However, current wireless sensor networks have problems in fire monitoring over wide areas because of limited battery capacity and the short life spans of the sensor nodes [1]. Unlike general environmental data monitoring such as precipitation or climate monitoring, fire monitoring requires very frequent real-time data transmissions until the fire is put out. Consequently, the sensor nodes in fire monitoring applications consume much more energy in relatively short periods of time than do sensor nodes in general environmental data monitoring applications, and they finally have shorter life spans [2].

For fire monitoring applications, flat-based routing protocols for sensor networks, such as SPIN (Sensor Protocols for Information via Negotiation), are not appropriate since all the sensor nodes that detect fire start sending fire data to the sink node individually and consume energy of the sensor network very quickly. The existing cluster-based routing protocols for sensor networks, such as LEACH (Low-Energy Adaptive Clustering Hierarchy) and TEEN (Threshold-sensitive Energy Efficient sensor Network protocol), are more energy-efficient for fire monitoring applications than the flat-based routing protocols since each cluster head can collect fire data from local sensor nodes, encapsulate the collected data into a single data packet, and transmit the packet to the sink node at a time. However, as the number of clusters that detect fire increases due to the wide propagation of the fire, the energy efficiency of fire monitoring with cluster-based routing protocols will decrease [2–6].

This paper proposes EFMP (Energy-efficient Fire monitoring Protocol), a fire monitoring protocol operating over cluster-based sensor networks. To further increase the energy efficiency of cluster-based sensor networks for fire monitoring, EFMP reduces the number of transmissions of fire data from the cluster heads to the sink node by dynamically creating and reorganizing the sensor network cluster hierarchy

according to the fire propagation over the sensor network clusters. Fire monitoring experiments showed that EFMP consumes 41% and 12% lesser energy on average per node than do LEACH and TEEN, respectively. Furthermore, the number of sensor nodes that survived in the EFMP-based experiment was 12% and 14% more than that in the LEACH- and TEEN-based experiments, respectively.

The rest of this paper is organized as follows. Section 2 analyzes the existing energy-efficient sensor network protocols and the problems faced when using these protocols in fire monitoring applications. Section 3 describes the design of the EFMP proposed in this study. Section 4 compares the performance of EFMP in fire monitoring environments with the performances of existing cluster-based sensor network protocols using NS-2. Finally, Section 5 summarizes the paper and future directions as conclusion.

2. Related Works

This section explains typical energy-efficient routing protocols for sensor networks such as SPIN, LEACH, and TEEN, which can be used in fire monitoring applications.

2.1. SPIN. SPIN is a method proposed for avoiding repetition of data transmission when the same data are sent to multiple nodes. SPIN sends ADV (Advertisement) messages from sensors collecting data. It asks for readiness to receive data by nearby sensors and the sensor receiving the ADV message sends a REQ (Request) message once it is ready. The sensor receiving the REQ message sends its own sensor data. However, the REQ is not sent if it is not ready. Data are, therefore, sent only to sensors that sent REQs since sensors not sending REQs are not ready. The drawback of SPIN in fire monitoring system environments is that the sensor only sends a data packet once it receives a REQ after sending an ADV packet. If a sensor receives four REQ packets, then the sensor sends four data packets individually. Packets are delivered to nodes that are not relevant to the actual transmission path in this case and this wastes battery power [2, 3]. The main drawback is high consumption of battery power by numerous transmissions from networks requiring many nodes, as in the case of fire monitoring system environments, since the sensors are connected un hierarchically. In addition, transmission is not completed if there are isolated sensors, and battery power is consumed between nearby nodes and isolated nodes since ADV packets continue to be sent for connection. These problems may occur in the case of a forest fire where a large number of sensors are installed over wide area [7–9].

2.2. LEACH. LEACH is a protocol for forming efficient clustering; sensor nodes form local clusters by themselves and distribute energy evenly. One node-forming cluster acts as a cluster head, and the sensor node with the most energy operates as a cluster head; when its energy capacity becomes smaller than those of the other nodes, another node with the most energy then performs the role of cluster head. The battery drains quickly when a sensor node acts as a cluster head; the burden on the node operating as the cluster head is, therefore, reduced by rotating the cluster heads. LEACH is currently on of the most popular method for environmental monitoring system. However, LEACH may waste a lot of energy since each individual cluster head operates independently in a fire monitoring system with a large number of sensor nodes. The number of related clusters increases as the fire spreads widely. Increases in the number of cluster heads sending information and the number of transmissions in the sensor network will result in a reduction in energy efficiency [2, 3, 10, 11].

2.3. TEEN. TEEN is a cluster-based and reactive routing protocol that works with two threshold values: a hard threshold and a soft threshold. A hard threshold is a threshold value for the sensed attribute. Nodes sensing this value only turn their transceivers on if the sensed value is above the defined threshold. A soft threshold is a small change in the value of the sensed value; it triggers the node to switch on its transceiver to become active and transmit. The sensed value is stored as an internal variable in the node. Every time a new cluster head is selected, the threshold values can change. When the cluster node's value exceeds the soft threshold, then it starts sensing again. The advantage of this protocol is the reduced number of deliveries, making it highly efficient in terms of energy consumption and response time [3, 4, 12, 13]. Nevertheless, TEEN may also suffer from the drawback of typical cluster-based routing protocols in fire monitoring environments.

3. Design of EFMP

3.1. Protocol Stack for EFMP. The method of transmission from the sensor to the cluster head and the cluster head to the sink is based on the existing cluster structure. That is, the methods for partitioning the sensor network into clusters and election of cluster heads are done by the underlying cluster-based routing protocol.

Figure 1 shows the EFMP protocol stack designed in this study. The stack is broadly composed of the system layer and the protocol layer. The system layer includes the sensor node OS and hardware, and the protocol layer includes the MAC (Medium Access Control) protocol responsible for communication by the sensors. The protocol layer has cluster-based routing protocol, such as LEACH and TEEN, over the MAC protocol. The EFMP proposed in this paper is on top of the protocol stack and can improve energy efficiency by reducing the number of transmissions by dynamic hierarchical clustering, while maintaining the existing protocols; this is the main function of the EFMP since it controls cluster hierarchy by operating on existing cluster-based routing protocols.

3.2. Cluster Hierarchy and Roles of Cluster Heads. The major difference between the sensor network cluster organization of an EFMP and that of existing cluster-based routing protocols is that the EFMP has a hierarchical cluster structure that can be reorganized dynamically. As illustrated in Figure 2, in an EFMP, the cluster heads are layered and classified

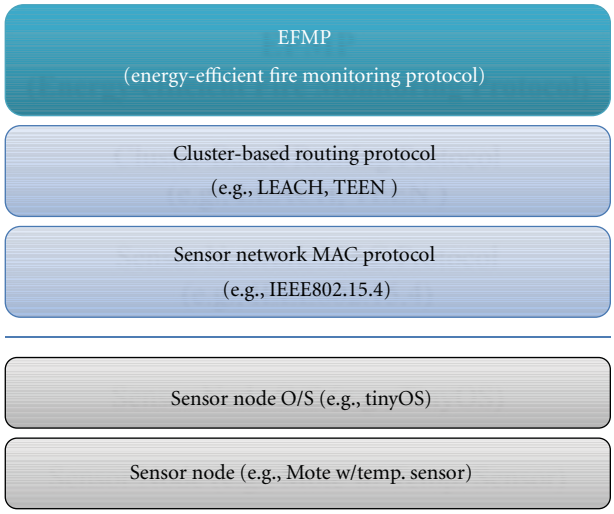


Figure 1: EFMP protocol stack.

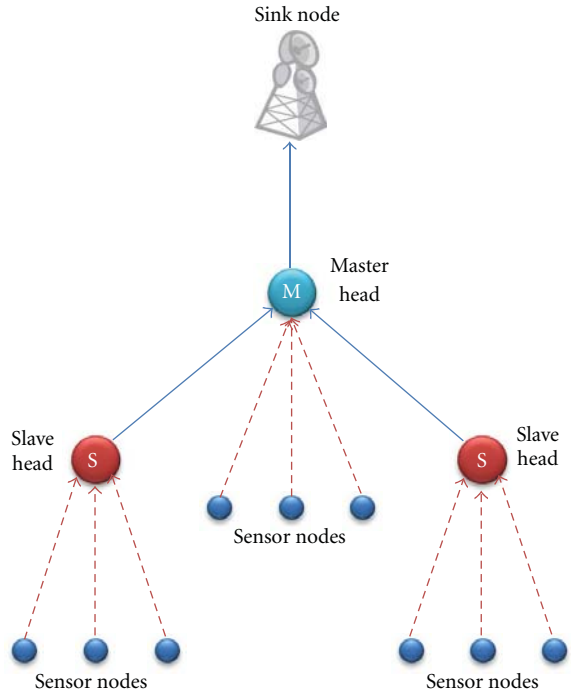


Figure 2: Cluster hierarchy of EFMP.

into master heads and slave heads. The structural difference between a general cluster structure and the proposed EFMP general cluster structure is the presence of a master head that collects and manages data from cluster heads. The master head collects information from the cluster heads and sends the collected information to the sink node.

Existing clusters only collect and send information from the cluster head. They use up large amounts of energy since each cluster head independently transmits data if there are many cluster heads. EFMP achieves better efficiency by reducing the number of transmissions compared with cluster-based cluster head methods by electing a master head and transmitting information by collecting data, in the case of a fire, from sensor nodes, according to information on the fire, as shown in Figure 2. Sensor nodes are the smallest unit forming a sensor network. All the sensor nodes can be slave heads or master heads, depending on the environment. If a master head is elected among the cluster heads, all the cluster heads except the master head become slave heads.

Slave heads send information to the master head. The master head sends information to the sink node in batches by collecting data received from the slave heads. From among the existing master head candidates, the cluster head with the least number of transmissions to the sink node is selected as the master head. If the number of transmissions is the same, then the one that is closest to the sink node, with the least number of sensors detecting fire, or the cluster head with the most battery power, is selected as the master head.

Figure 3 shows the changes in the roles of the nodes according to changes, in the case of a fire, and all the nodes are initialized while in watch mode.

Watch mode refers to the initial state, in which the sensors did not detect a fire. The cluster head that first detects a fire changes from watch mode to master mode. It transforms the nearby cluster heads in watch mode to slave mode since it is in master mode itself, that is, the EFMP system displays much better energy efficiency than existing cluster networks do since the system makes hierarchical networks dynamic by changing the cluster head into a cluster head in watch mode, slave mode, or master mode, according to the direction of the fire.

3.3. Types of Packets. The features of EFMP are electing a master head that collects and manages information from slave heads managing clusters. The location of the master head changes according to information on the fire (direction of the fire). The passage defines the structural design of the EFMP and the algorithm and packet type for deciding a master head.

(1) SIG_FIRE Packet. If a sensor node detects fire or fire data, it immediately sends a SIG_FIRE packet to its cluster head. The fire is detected if the following condition is met, where TEMP (t_n) represents the temperature measured by a sensor node at the current time t_n, $(\sum_{t=t_1}^{t_{n-1}} \mathrm{TEMP}(t))/(n-1)$, represents the average temperature measured from time t_1 to t_{n-1} (i.e., the average of all temperatures measured before t_n):

$$\left| \mathrm{TEMP}(t_n) - \frac{\left(\sum_{t=t_1}^{t_{n-1}} \mathrm{TEMP}(t)\right)}{(n-1)} \right| > \Delta\mathrm{TEMP}_{\mathrm{MAX}}. \quad (1)$$

Thus, if the difference between the currently measured temperature and the average of the previously measured temperatures is greater than a specific limit $\Delta\mathrm{TEMP}_{\mathrm{MAX}}$, the sensor node transmits a SIG_FIRE packet to its cluster head.

Figure 4 shows the structure of the SIG_FIRE packet. As shown in the figure, the SIG_FIRE packet consists of Sensor_ID and Sensor_data fields. The Sensor_ID field includes

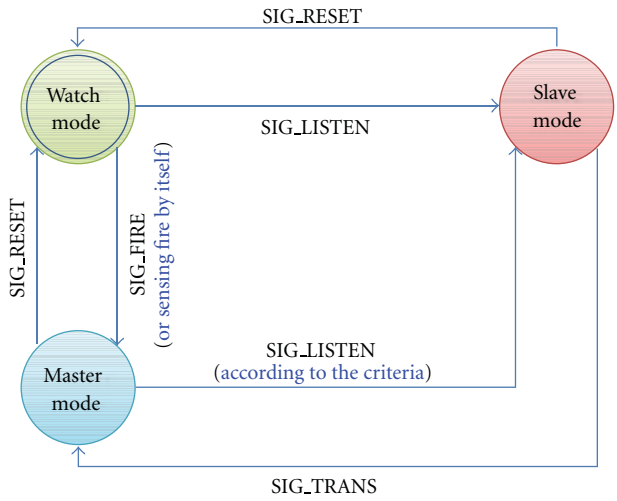


Figure 3: Role-transition diagram for cluster heads.

Type	Content	
SIG_FIRE	sensor_ID	sensor_data

Figure 4: SIG_FIRE packet format.

the unique identifier of a sensor node (e.g., the MAC address or IP address of a Zigbee node), and the Sensor_data field includes currently measured temperature data.

(2) SIG_DATA Packet. SIG_DATA is gathered at each cluster head to be transmitted to the sink node. The sensor monitors the fire and sends information to the slave heads by detecting fire data. The sensor node receiving information relays the fire data (SIG_FIRE) collected to a master head, and SIG_DATA is used when sending information from slave heads to the master head. Figure 5 shows the structure of the SIG_DATA packet. The SIG_DATA packet consists of Head_ID, Sensor_ID_N, and Sensor_data_N fields. The Head_ID field is a unique identifier of the cluster ID and the Sensor_ID_N is the sensor node's ID in the cluster header. Sensor_data_N indicates the fire data collected from all the sensor nodes in the cluster head.

(3) SIG_INFORM Packet. The SIG_INFORM packet is used when a new master head provides its own information to a slave head. It is used when the first slave head within a cluster detects a fire and is changed to the master head and provides master head information to nearby sensors as well as information about the new master head when the master head is replaced. Figure 6 shows structures of SIG_INFORM packet. The SIG_INFORM packet consists of Head_ID, Bat_capacity, Num_sensors, Sensor_ID_N, and Sensor_data_N fields.

The Head_ID field is a unique identifier of the cluster ID and Bat_capacity is the remaining battery capacity. Num_sensor is the number of sensor nodes, and Sensor_ID_N is the sensor node's ID in the cluster head. Sensor_data_N is the fire data collected from all the sensor nodes in the cluster head.

(4) SIG_QUERY and SIG_RESP Packets. The slave head within the cluster is transformed to the master head once the sensor detects a fire within its own cluster. The slave head becomes a master head candidate and the previous master head asks nearby master head candidates whether it is the appropriate master head. The packet format sending this information is called SIG_QUERY and the response is called SIG_RESP. Figures 7 and 8 show the structures of the SIG_QUERY packet and the SIG_RESP packet. The SIG_QUERY packet has a unique identifier, Master_ID, and Num_candidate is the number of master candidates. SIG_RESP has a unique identifier, Slave_ID. The Hop_to_master field is the number of hops between the master heads. Candidate_ID_N is the master head candidate.

(5) SIG_LISTEN Packet. Address information on the new master head should be sent to nearby slave heads once a new master head has been elected. Figure 9 shows SIG_LISTEN packet. The SIG_LISTEN packet provides the address information of the newly elected master head.

(6) SIG_TRANS and SIG_RESET Packets. Figure 10 shows structures of the SIG_TEANS, SIG_RESET packet. SIG_TRANS is used for registering information on the new master head by receiving the SIG_LISTEN packet from the new master head. It is also used when a slave head registers its own address information. The SIG_RESET packet is used when the previous master head again becomes a slave head by handing over its authority as a master head to the newly elected master head, and when the reverted previous master head registers its information with the new master head.

3.4. EFMP Operating Procedures. EFMP not only sends information acquired by fire detection but also provides hierarchical clusters by selecting a master head according to the direction of the fire and reduces energy consumption by sensors. EFMP is composed of a FIRE_Detection part for detecting a fire and a Startup_Monitoring part for sending information on fires. This section describes the detailed operating procedures of EFMP in two parts.

(1) Fire Detection in Sensor Nodes. Fire detection procedures refer to sending information to cluster heads by detecting fire for the first time from the information collected by sensors. A sensor node has an initial value of 0 for TEMP_{SUM}, the current temperature as the initial value of TEMP_{VEA}, and 1 for the initial value of n, as shown in Figure 11. The sensor node sends the SIG_FIRE packet to the cluster head if the current temperature value is inserted into TEMP_{CUR} and $|\text{TEMP}_{\text{CUR}} - \text{TEMP}_{\text{AVE}}|$ is greater than $\Delta\text{TEMP}_{\text{MAX}}$.

(2) Fire Monitoring Startup. The cluster head among the sensor nodes receiving the SIG_FIRE packet becomes aware

Type	Content								
SIG_DATA	head_ID	num_sensors	sensor_ID_1	sensor_data_1	sensor_ID_2	sensor_data_2	...	sensor_ID_*N*	sensor_data_*N*

Figure 5: SIG_DATA packet format.

Type	Content							
SIG_INFORM	head_ID	bat_capacity	num_sensors	sensor_ID_1	sensor_data_1	...	sensor_ID_*N*	sensor_data_*N*

Figure 6: SIG_INFORM packet format.

Type	Content					
SIG_QUERY	master_ID	num_candidate	candidate_ID_1	candidate_ID_2	...	candidate_ID_*N*

Figure 7: SIG_QUERY packet format.

Type	Content					
SIG_RESP	slave_ID	hops_to_master	num_Candidate	hops_to_Candidate_1	...	hops_to_Candidate_*N*

Figure 8: SIG_RESP packet format.

Type	Content
SIG_LISTEN	master_addr

Figure 9: SIG_LISTEN packet format.

Type	Type
SIG_TRANS	SIG_RESET

Figure 10: SIG_TRANS and SIG_RESET packet format.

of a fire breaking out from its own cluster. Figure 12 shows procedure of fire monitoring startup The cluster head recognizing the fire transforms to master mode from watch mode, and it sends a SIG_LISTEN packet containing detection of fire to the other cluster heads and sends the SIG_DATA packet to the sink node.

(3) Monitoring Procedure. Figure 13 shows monitoring procedure. The sensor node for detecting fire has the function of detecting fire continuously. Detected information is sent to slave heads and the master head through a SIG_FIRE packet, and each slave head collects and relays sensor information. The slave head in a fire zone, on detecting a fire within its own cluster of slave heads, transmits that it is a candidate for master in the case of a fire through the SIG_INFORM packet. Conversely, slave heads send the SIG_DATA packet to the master head, thinking that they are not in a fire zone when they do not detect a fire. The master head sends all the collected SIG_DATA received from the multiple slave heads to the sink node in the batch.

(4) Master Election. The master head of an EFMP is not a fixed type but can change continuously, depending on the fire. In other words, a new cluster will participate in fire monitoring when the fire spreads to a new cluster head. The following procedures are used to elect the new master head.

Suppose that M represents a set of master candidates that includes the current master head and all the other heads newly involved in fire monitoring, and the element of M are represented as $m_1, m_2, \ldots, m_L$. In addition, existing slave heads are represented by $s_1, s_2, \ldots, s_N$ and the sink node is represented by sink. The master election criteria are represented by the following formula, where DIST (a, b) represents the number of transmissions from node a to node b:

$$\begin{aligned} \mathrm{DIST}_{\mathrm{tot}}(m_x) = {} & \sum_{i=1}^{N} \mathrm{DIST}(s_i, m_x) \\ & + \left(\sum_{j=1}^{L\ (\mathrm{excl.}\ x)} \mathrm{DIST}\left(m_j, m_x\right) \right) \\ & + \mathrm{DIST}(m_x, \mathrm{sink}). \end{aligned} \tag{2}$$

$\mathrm{DIST}_{\mathrm{tot}}\ (m_x)$ represents the total number of transmissions from each slave head to the sink node via the newly elected

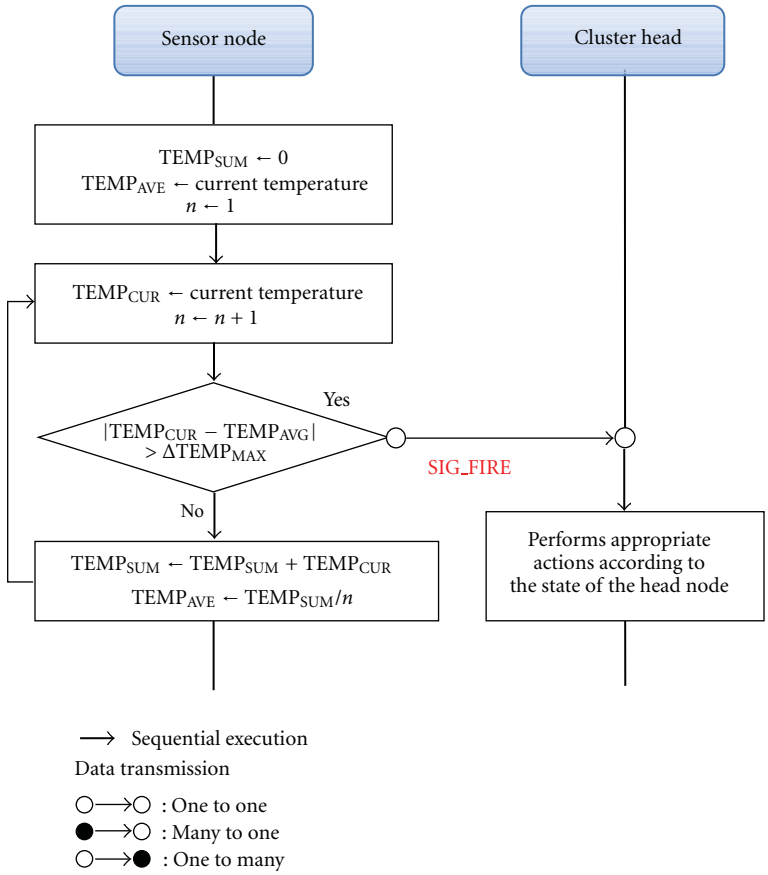


Figure 11: Fire detection procedure.

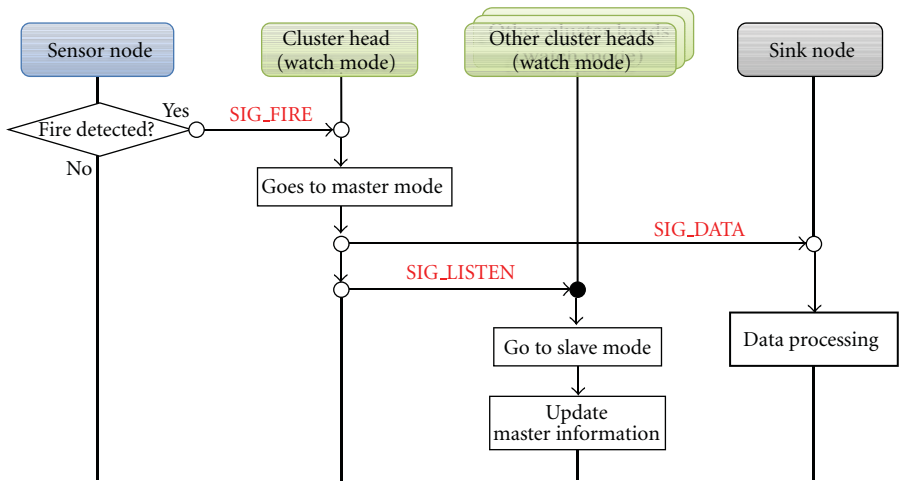


Figure 12: Fire monitoring startup procedure.

master head m_x. Thus, for all $m_x \in M$, the m_x that has the minimum $\mathrm{DIST}_{\mathrm{tot}}(m_x)$ value is elected as the new master head.

The current master head asks each master head candidate, using the SIG_QUERY packet, for the information needed to evaluate the master election criteria, and each master head candidate replies with the SIG_RESP packet, as shown in Figure 14. If a new master head is elected according to the master head criteria, the current master head notifies the new master head with the SIG_TRANS packet, and then, the new master head sends the other slave heads information about itself through the SIG_LISTEN packet.

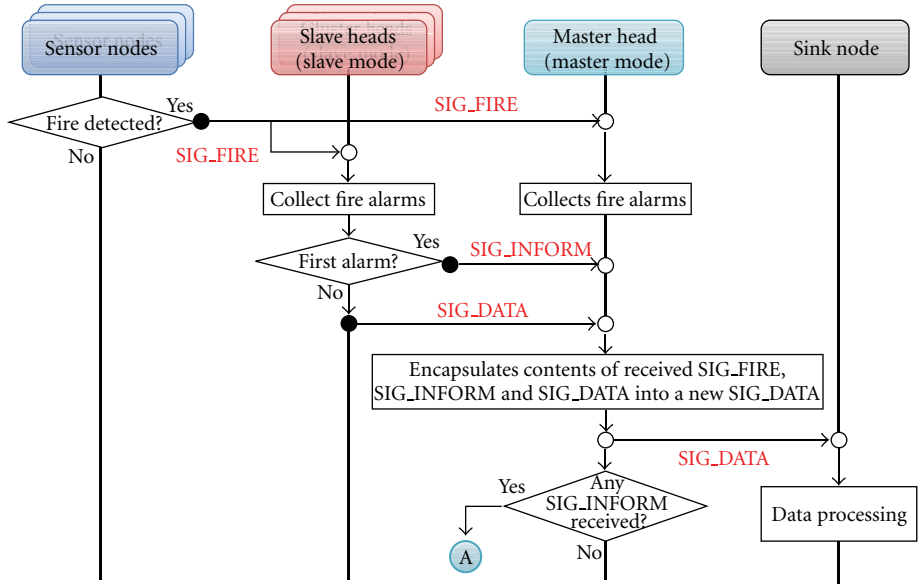


Figure 13: Monitoring procedure.

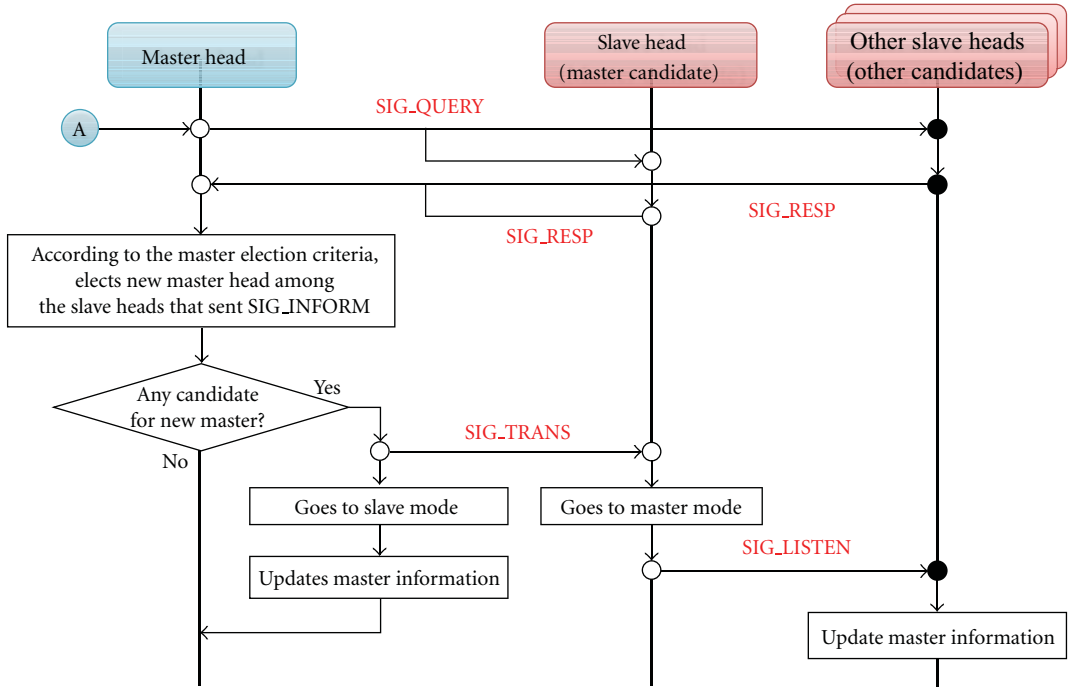


Figure 14: Master election procedure.

3.5. Fire Monitoring with EFMP. The cluster head in EFMP operates in watch mode before detection of a fire, monitors fires from its own zone, and collects sensor information. The sensor of a cluster head in watch mode, as shown in Figure 15(a), sends fire information to its own cluster head once a fire is detected within its own cluster.

The cluster head receiving information on a fire from the sensor recognizes a fire within its own cluster for the first time and changes into a master head from watch mode. First, the elected master head sends information on the fire within its own cluster to the sink node (Figure 15(b)). It then sends the information about itself to nearby cluster heads in the slave mode through a SIG_LISTEN packet (Figure 15(c)) and the slave heads update their master head information. The cluster head where the fire broke out is elected as the first master head. A new master head, elected according to the progress of the fire, sends collected data to the sink node, as shown in Figure 15(d). It acts as the master head until another appropriate master head candidate is elected.

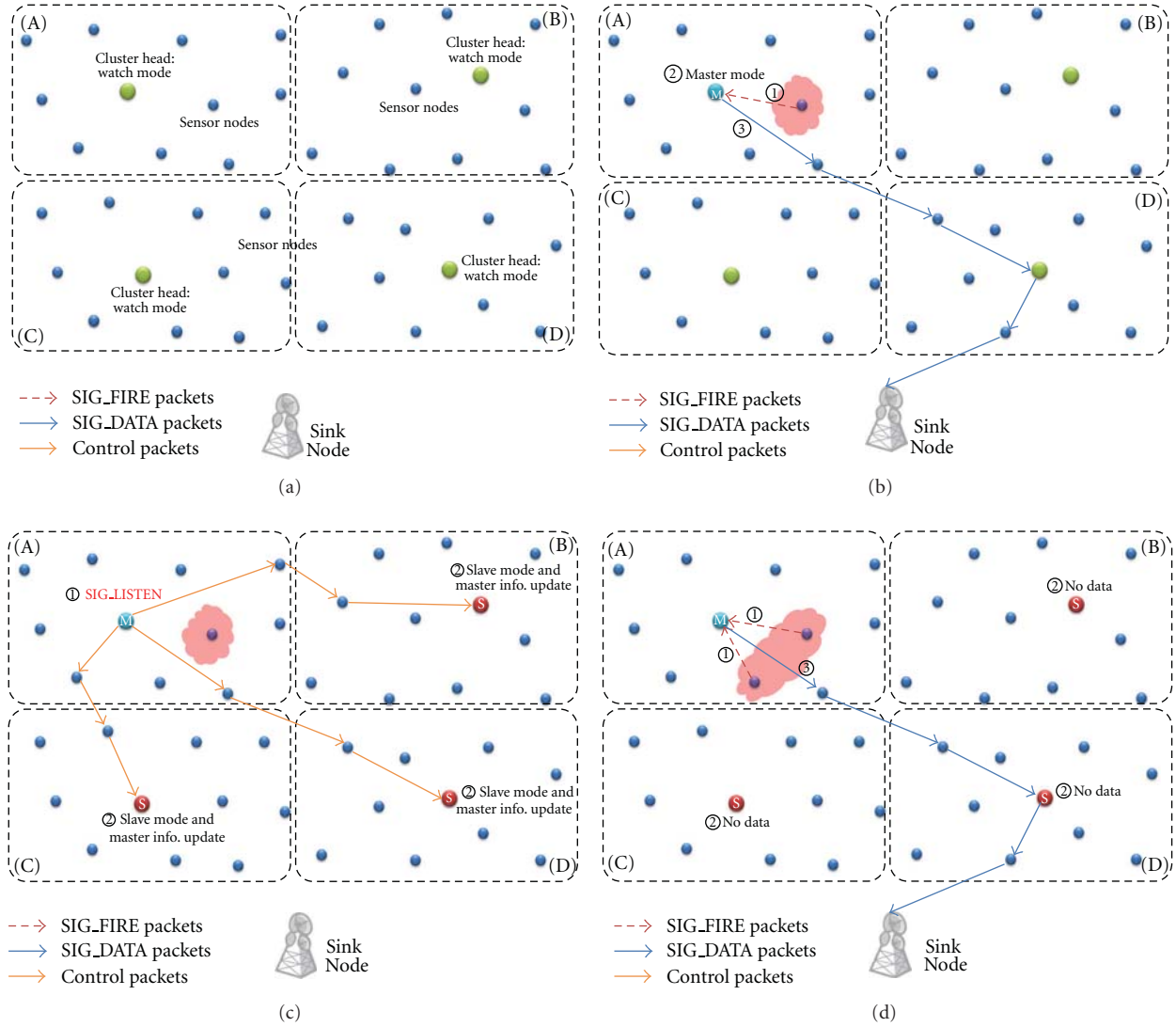


Figure 15: Method of electing master head from watch mode.

The number of cluster heads detecting fire within their own cluster increases as the fire spreads (Figure 16(a)). The cluster head detecting the fire can become a master head candidate, and it uses a SIG_INFORM packet to notify the current master head that it is a new master head candidate.

Figure 16(b) shows that the master head receiving the SIG_INFORM message sends each master head candidate a SIG_QUERY message asking for the total number of hops, and the master head candidates calculate the total number of hops to the sink and send the information to the master through a SIG_RESP message. The master head receiving the information compares the total number of hops for sending messages to the sink and the number of hops from each master head. It then asks the candidates having smaller numbers of hops than itself if they are available as the next master head, sends a message that it will maintain the role of the current master if the total numbers of hops from the master head candidates are greater, and maintains the sensor network. As shown in Figure 16, the master head changes from cluster head (A) to cluster head (B) since the fire spreads more widely than in (C). When it becomes the master head, cluster head b sends a SIG_LISTEN packet requesting updating of the new master head information to each master head candidate, and the master head candidates receiving the messages update the master head information.

The previous master head updates the information on the next master head itself but does not update after receiving SIG_LISTEN since it knows which one is the next master head; it sends a SIG_TRANS packet indicating completion of the update and thus elects a new master head. The elected master head relays information to the sink node in batches on receiving information from each cluster head, as the previous master head had been doing.

Each cluster head sends information collected from its own sensors to the sink node in the previous cluster head method, but there are problems with different energy efficiencies in transmission by cluster according to the location of the cluster head, so the location of the cluster head needs

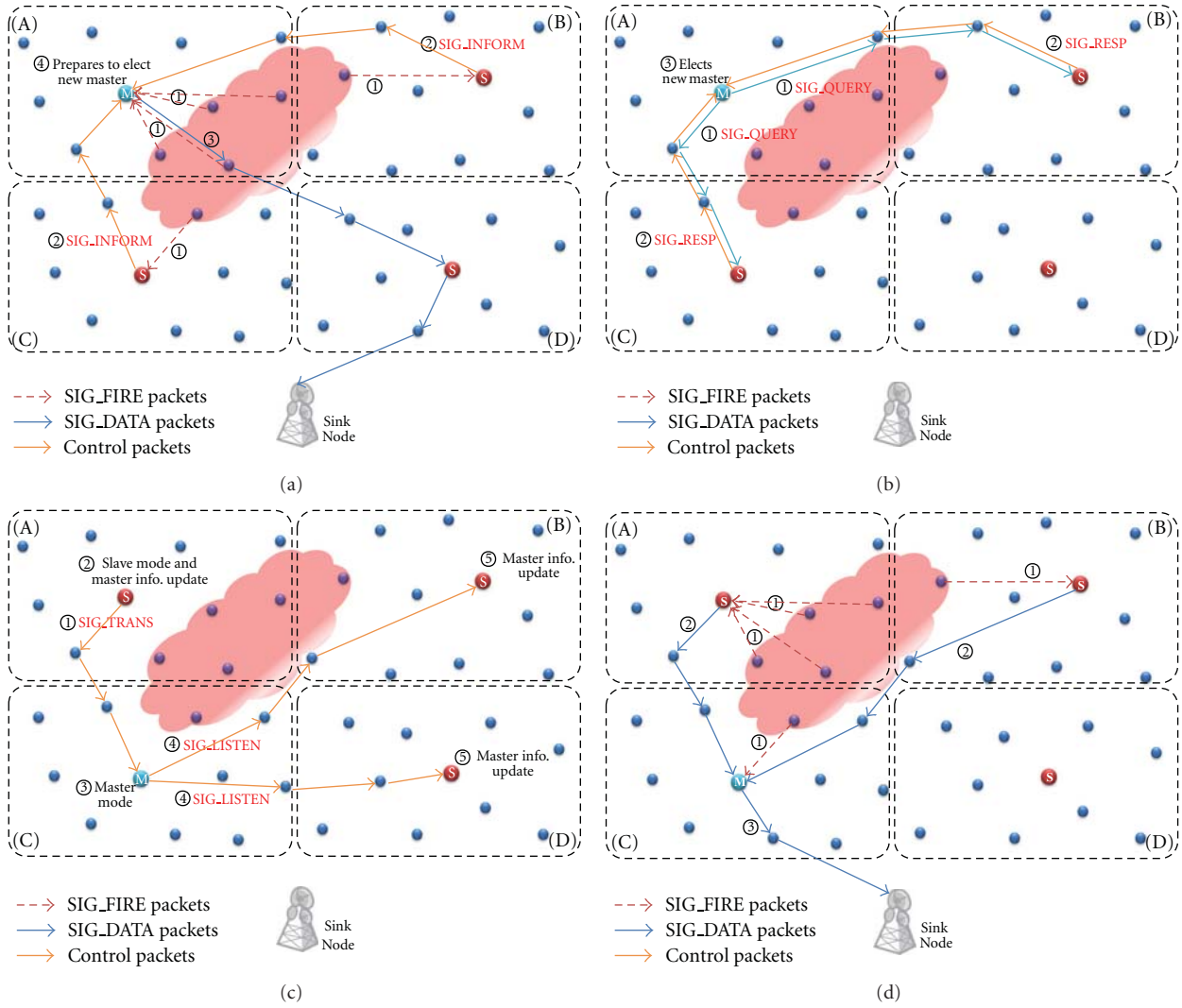


Figure 16: Method of electing new master head.

to be regularly changed to solve this problem. Otherwise, the overall energy efficiency of the sensor network worsens because the locations of several cluster heads are wrong. However, EFMP elects the optimal master head according to changes in the surroundings, and thus solves the problems of the cluster head method and increases the energy efficiency accordingly.

4. Experiments and Performance Evaluation

4.1. Experimental Environment. NS-2 [14] was used to evaluate the performance of EFMP in this study. A Mannasim module was used to implement the sensor network. The goal of Mannasim is to develop a detailed simulation framework that can accurately model different sensor nodes and applications, while providing a versatile testbed for algorithms and protocols. NS-2 supports modeling and simulations such as mobile and fixed wireless networks, wireless sensor networks, body area networks, ad-hoc networks, and vehicular networks for Mannasim [14, 15]. This study simulates a wireless sensor network using IEEE 802.15.4 (Zigbee) MAC. For our experimental environment, we simulated a fire monitoring sensor network by first placing 300 nodes that are clustered into 10 clusters consisting of 30 nodes per cluster in a bounding area of $100 \times 100\,\text{m}^2$. All the nodes start with an initial energy of 10 J. The performance evaluation compares LEACH with a LEACH-based EFMP (LEACH_EFMP) and TEEN with a TEEN-based EFMP (TEEN_EFMP) in terms of the number of surviving nodes and average power consumption per sensor node of the sensor network. Each round of experiments was performed for 180 min and a total of 10 rounds of experiments were performed under the same environment. It was assumed that the fire was propagated to one adjacent cluster at a time and the direction of propagation was random.

4.2. Experimental Results. Figure 17 shows the number of surviving nodes during the experiments. The proposed protocols, LEACH_EFMP and TEEN_EFMP, had larger numbers of nodes alive than did the original LEACH and TEEN

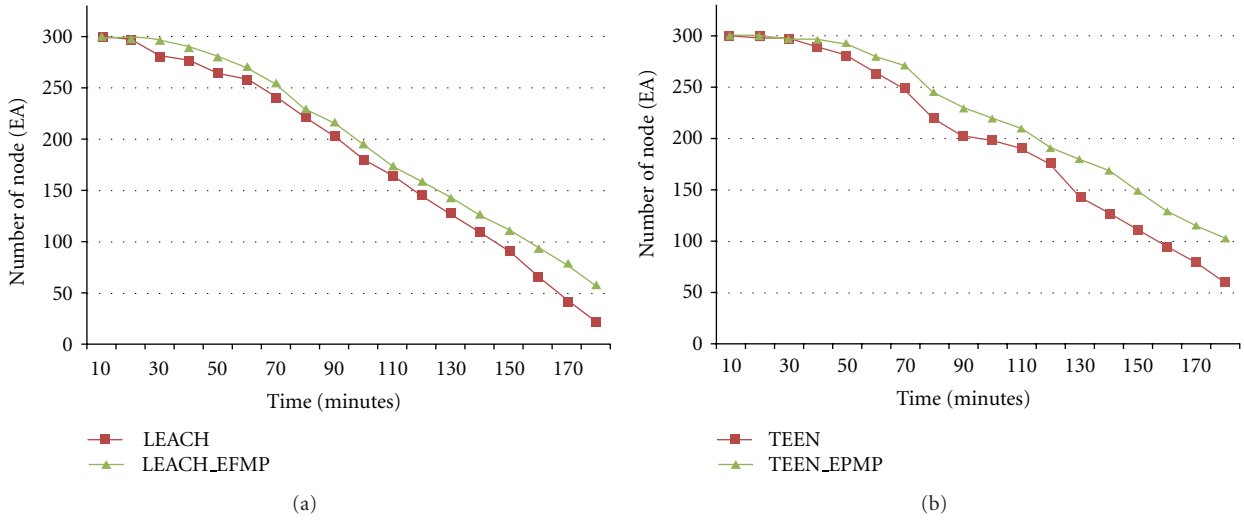


Figure 17: Number of surviving nodes.

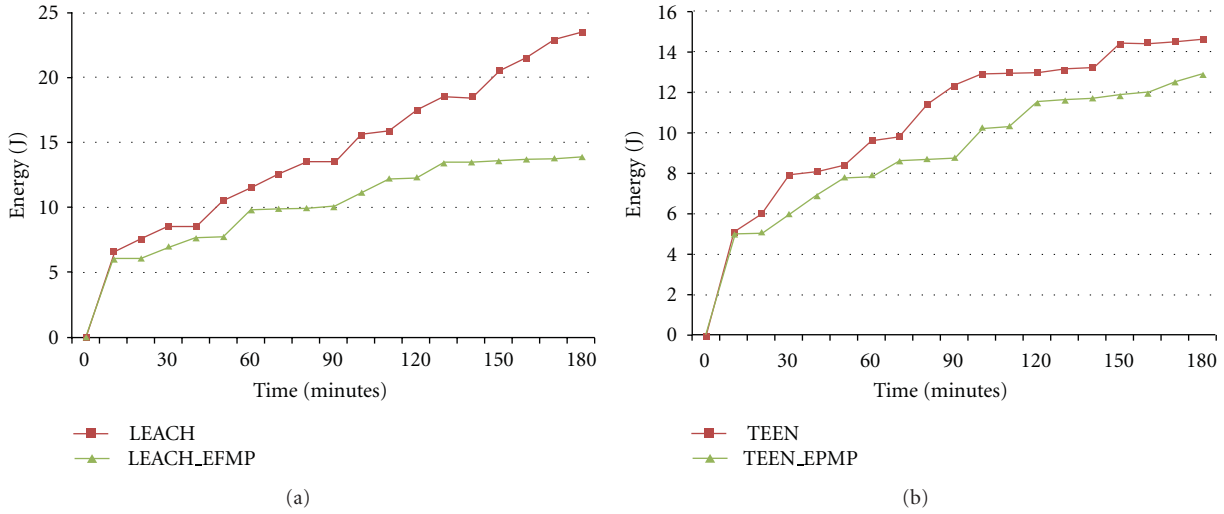


Figure 18: Average energy consumption over all sensor nodes.

protocols. When using LEACH_EFMP, 58 out of 300 nodes were still alive (i.e., approx. 19% of the entire nodes) after 180 min. However, only 22 nodes (i.e., 7% of the entire nodes) were alive when using the original LEACH protocol under the same condition. TEEN_EFMP also gave better results than did the original TEEN. TEEN_EFMP had 103 nodes (i.e., 34%) alive, as opposed to the case of the original TEEN protocol, where only 60 nodes (i.e., 20%) were alive.

Figure 18 shows the average energy consumption (including energy consumed for initial cluster formation of the underlying routing protocol) over all the sensor nodes during the experiments; this was calculated using the following formula, where $C(S_i, t)$ represents the energy consumed by a sensor node S_i at time t, and N represents the total number of sensor nodes:

$$\frac{\sum_{i=1}^{N} C(S_i, t)}{N}. \quad (3)$$

As shown in Figure 18, the average energy consumptions at the end of the experiments with LEACH and LEACH_EFMP were 23.5 J and 13.9 J respectively. The simulation results show that LEACH_EFMP used 41% lesser energy per sensor node on average than did LEACH under the same fire monitoring condition. The average energy consumptions of TEEN and TEEN_EFMP were 14.6 J and 12.9 J, respectively. The simulation results show that TEEN_EFMP used 12% lesser energy than did TEEN.

5. Conclusion

Sensor networks are the most appropriate systems in fire monitoring environments. However, sensor networks have problems in fire monitoring environments because of their limited battery capacity. The energy efficiency of a typical cluster-based sensor network drops when a large number of

sensor clusters simultaneously monitor a fire over a wide area, because of the fire monitoring characteristics. In this study, we proposed the EFMP that reduces overall energy consumption of the sensor network by dynamically forming a multilayer cluster hierarchy based on the propagation of the fire and efficiently transmitting data over the hierarchical cluster-based sensor network appropriate for fire monitoring.

The performance evaluation results showed that EFMP-based fire monitoring consumes 41% lesser energy on average per sensor node than does LEACH-based fire monitoring and 12% lesser energy on average per node than does TEEN-based fire monitoring. Furthermore, the number of sensor nodes surviving in the EFMP-based experiments was 12% and 14% more than that in the LEACH- and TEEN-based experiments, respectively.

Finally, future work will focus on estimating the size, speed, and direction of a fire by extending EFMP in general fire monitoring environments. We hope that this study will contribute to protecting more lives and properties from fire disasters.

Acknowledgment

This paper was supported by Konkuk University in 2010.

References

[1] 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*, vol. 8, pp. 1–10, January 2000.

[2] F. Xiangning and Y. Song, "Improvement on LEACH protocol of wireless sensor network," in *Proceedings of the International Conference on Sensor Technologies and Applications (SENSORCOMM '07)*, pp. 260–264, October 2007.

[3] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–105, 2002.

[4] A. Manjeshwar and D. P. Agrawal, "TEEN: a routing protocol for enhanced efficiency in wireless sensor networks," in *Proceedings of the 15th International Parallel and Distributed Processing Symposium (IPDPS '01)*, vol. 3, pp. 2009–2015, 2011.

[5] M. J. Handy, M. Haase, and D. Timmermann, "Low energy adaptive clustering hierarchy with deterministic cluster-head selection," in *Proceedings of the 4th IEEE Conference on Mobile and Wireless Communications Networks*, pp. 368–372, 2002.

[6] L. Zou, M. Lu, and Z. Xiong, "A distributed algorithm for the dead end problem of location based routing in sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 4, pp. 1509–1522, 2005.

[7] D. E. Heim, R. E. Fontana, C. Tsang, V. S. Speriosu, B. A. Gurney, and M. L. Williams, "Design and operation of spin valve sensors," *IEEE Transactions on Magnetics*, vol. 30, no. 2, pp. 316–321, 1994.

[8] J. Qiangfeng and D. Manivannan, "Routing protocols for sensor networks," in *Proceedings of the 1st IEEE Consumer Communications and Networking Conference (CCNC '04)*, pp. 93–98, January 2004.

[9] J. N. Al-Karaki and A. E. Kamal, "Routing techniques in wireless sensor networks: a survey," *IEEE Wireless Communications*, vol. 11, no. 6, pp. 6–28, 2004.

[10] A. A. Ahmed and H. Shi, "A survey on network protocols for wireless sensor networks," in *Proceedings of the of Information Technology: Research and Education (ITRE '03)*, pp. 301–305, 2003.

[11] S. D. Muruganathan, D. C. F. Ma, R. I. Bhasin, and A. O. Fapojuwo, "A centralized energy-efficient routing protocol for wireless sensor networks," *IEEE Communications Magazine*, vol. 43, no. 3, pp. S8–S13, 2005.

[12] C. Patra, A. Guha Roy, S. Chattopadhyay, and P. Bhaumik, "Designing energy-efficient topologies for wireless sensor network: neural approach," *International Journal of Distributed Sensor Networks*, vol. 2010, Article ID 216716, 7 pages, 2010.

[13] Y. Gu, Q. Wu, and N. S. V. Rao, "Optimizing cluster heads for energy efficiency in large-scale heterogeneous wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 2010, Article ID 961591, 9 pages, 2010.

[14] "Network Simulator-2," http://isi.edu/nsnam/ns/.

[15] "Mannasim for Wireless Sensor Network Simulation," http://www.mannasim.dcc.ufmg.br/.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 506810, 9 pages
doi:10.1155/2012/506810

Research Article

An Adaptive System Supporting Collaborative Learning Based on a Location-Based Social Network and Semantic User Modeling

Sook-Young Choi[1] and Jang-Mook Kang[2]

[1] *Department of Computer Education, Woosuk University, Jeollabukdo 565-701, Republic of Korea*
[2] *Electronic Commerce Research Institute of Dongguk University, Gyeongsangbuk-do 780-714, Republic of Korea*

Correspondence should be addressed to Jang-Mook Kang, mooknc@gmail.com

Received 2 November 2011; Accepted 2 January 2012

Academic Editor: Tai Hoon Kim

This paper presents an adaptive e-learning system which supports collaborative learning based on a location-based social network and semantic modeling. In the system, a social network among e-learning learners is dynamically constructed on the basis of the location information of learners using GPS sensors for collaborative learning. In addition, user modeling for semantic social network and adaptive e-learning is supported. For these, a mechanism that supports the construction of a dynamic social network using the location information of smart phones is provided. The user modeling for the construction of a semantic social network and the support of adaptive e-learning considers the information such as the users' achievement, preference, and learning goal. Through this system, e-learners can find mentors who are able to facilitate their learning and create communities for collaborative learning offline and online. The support of a location-based social network service and semantic user model in our system would increase interactions among e-learners and improve satisfaction regarding their mobile learning environment.

1. Introduction

E-learning offers higher education the opportunity to expand the borders of classrooms to include distance learners. E-learning gives an attractive learning opportunity for learners who are restricted by time and space, thus, increasing the number of e-learners. However, there is an issue about the high drop-out rate associated with online courses. While some e-learners thrive on the increased flexibility that the medium provides, others languish in isolation and struggle to get started [1, 2]. Therefore, as colleges continue to attract new online learners, administrators are also trying to find ways to keep them enrolled.

Tinto [3] stresses that academic satisfaction is not enough for some learners who suffer from isolation. The intensity and reciprocity of a social interaction can, together with other factors, result in such drastic measures as learners dropping out of a course. Carr [4] points out that anecdotal evidence and studies by individual institutions suggest that online course completion is much lower than in F2F (face to face) courses. A number of studies have found the retention of e-learners to be lower than the retention of oncampus learners. Interaction with classmates and the professor is a significant contributor to perceived learning in online courses [1, 5]. Learners who report a high level of interaction report a high level of perceived learning in a course.

The social dimension of learning is also central to the idea of situated learning. Social interaction has always been of great significance to teachers, learners, and others [3]. Learning is a function of the activity, context, and culture in which it occurs, where social interaction is critical. Accordingly, it would be important to boost e-learners so that they can construct a social network among themselves. Through the social network service, learners can have face-to-face meetings as well as online meetings. This social interaction would increase the learner's satisfaction with the course, increasing the probability that the learner will not drop the course [6].

Mobile devices can facilitate social interaction and access to information resources anytime and anywhere [7]. With proper design, applications in mobile devices can also facilitate learning. Currently, there are increasing demands and

interest in location-sensing based services with advancements in smart phones (which have GPS capability), PDAs, Bluetooth, dedicated GPS equipment, and other devices (such as i-Pad, navigation devices, and digital cameras). Mobile social networking is available to users as a consequence of social network services coming to mobile devices, especially smart phones [8].

However, the notion of modeling learners in a social network is very useful if applied to an e-learning system where the learners are physically in different locations and their social life is completely separated from academic life. These learners still need friends, who share the same interests, preferences, or learning experiences, to communicate and collaborate with each other. Therefore, user modeling is required to construct a semantic social network adapted to the special needs and personality characteristics of each learner so that the users' collaboration can be more efficient. That is, social networks are constructed on the basis of semantic user modeling.

This paper proposes a system that supports the construction of a social network service using the location information of the smart phone in mobile learning. This system also provides a mechanism to form a semantic social network among the learners who have similar learning interests, preferences, and learning experience based on user modeling information. Through this system, e-learners can create more effective communities for learning and exchanging help. That is, they can have face-to-face meetings as well as online meetings for collaborative learning. This support of a social network in our system based on location and semantic user modeling would increase interactions among e-learners and improve satisfaction regarding their e-learning learning environment. Consequently, it would make e-learning course completion rates higher.

2. Related Work

2.1. Advantages and Disadvantages to E-Learning. Knowing e-learning advantages and disadvantages helps learners use it effectively as well as select proper online programs for their learning. Thus, it is important to know the merits and demerits of e-learning. E-learning has advantages and disadvantages. Regarding the advantages of e-learning, we can consider following things.

E-learning reduces the learners' travel costs and time to and from school and they can study wherever they have access to a computer and Internet. It supports an adaptive learning to learners. That is, learners may have the option to select learning materials that meet their level of knowledge and interest and work at their own pace. In e-learning, different learning styles are addressed and facilitation of learning occurs through varied activities. In addition, it can improve the self-directed learning ability of learners. In other words, successfully completing online or computer-based courses builds self-knowledge and self-confidence and encourages learners to take responsibility for their learning [1, 4, 5].

The disadvantages of e-learning are as follows. The fact that a real person, to whom one might address questions and comments, may not be available is a concern. The programs do offer assistance to any learner that needs it, but the type of help may not be as helpful to learners if they are used to one-on-one and face-to-face assistance. This may be a little frustrating. In addition, unmotivated learners or those with poor study habits may fall behind. That is, learners who lack the ability for self-directed learning may face difficult challenges. Furthermore, learners may feel isolated or miss social interaction. Interaction between learners and their peers as well as their instructors is a significant aspect of e-learning. Thus, social interaction seems to be the common denominator among strategies and practices aimed at retaining online learners. To enhance retention of online learners, there should be supports for forming a social network between learners [1, 2, 5].

2.2. Social Network for Learning. Learning is a social network relationship. It is a shared or common experience as colleagues explore a new area and attend classes and lectures together, thereby, gaining a similar view of subject areas. The importance of interpersonal interaction in learning is undoubted. Several learning theories put special emphasis on the effects of interpersonal interaction on learning outcomes [9]. For example, collaborative learning theory assumes that learning emerges through interactions of an individual with others [7]. Constructivism regards learning as a social process that takes place through communication with others. The learner actively constructs knowledge by formulating ideas into words, and these ideas are built upon the reactions and responses of others [10]. Learning communities also give learners the opportunity to meet both social and academic needs simultaneously. Tinto [3] stresses that social affiliations serve as a vehicle through which academic involvement is engaged. This emphasis on the importance of the support provided by peers is seconded by Kinnunen and Malmi [11], among others.

In e-learning based on a distributed learning environment, a social network plays a more important role in support of the learners' learning. Through strengthening connections and inspiring communications among the learners, the learning of the whole community is promoted. Wegerif [12] highlights the importance of the social side of learning when designing a course, more specifically in an asynchronous learning network. Studies of social networks show that a social network exerts its effect on learning processes and effectiveness [12–14].

2.3. Social Networks and Location Sensing Information-Based Social Services. The provisioning of services using location information is known as location-based services (LBSs). There are more and more location-based experiences occurring in our daily lives such as location-based information services, location-based games, and location-based ubiquitous learning [10]. Mobile positioning is a technology of LBS, which can obtain the location of the mobile devices and their users. Mobile users cannot only query their positions and request services based on their current locations but also receive the information that they would be interested

in according to their current position [7]. For example, a passerby can query the information of the nearest coffee shop based on his current location, and the coffee shop can also send an advertisement to a passerby who is within a certain age range and near the shop. Recently, one of the emerging research topics is to utilize the location-awareness of mobile devices to further strengthen mobile-learning. Chen and Tsai [10] proposed a personalized context-aware ubiquitous English vocabulary learning system which can exploit appropriate context-awareness based on a learner's location. Mobile social networking appears to the users as the result of combining a social network service with smart phones. In relation to mobile social networking, a few studies have been proposed. Liao et al. [15] proposed a GPS-data-driven social networking service where people can share life experiences and connect to each other with their location histories. Li and Du [9] proposed a dynamic social networking system which supports location-based services. The system enables the participants with a common interest to communicate and share information with others within a certain geographical range in a decentralized mode.

However, these studies have not considered an e-learning environment. In addition, they have not proposed a practical method of how to use a location-based social network service in a realistic situation. However, our study considers a location-based social network service that effectively supports e-learning. Our system provides a mechanism to form social networks among the learners who take the same online course. Accordingly, e-learners can create communities for learning and have face-to-face meetings for collaborative learning, based on the location information.

2.4. Ontology-Based User Modeling for Social Networks. There are several approaches to representing and storing a user model in a web-based service system [16]. The most obvious is the use of a relational database to store data about the user. Using a relational database offers good performance and several other advantages such as security and data recovery. However, user models of web-based information systems often contain semistructured data as they use an overlay model, which follows the representation of the information space with various characteristics defined for concepts from the domain model. Relational databases are not primarily designed to express semistructured data. Moreover, relational databases are not well suited when frequent changes in data structure need to be performed, which is often the case in user modeling.

Another frequently used approach in the current web-based adaptive systems is the representation of the user model by an XML-based language using the file system that results in sufficiently powerful expressiveness. Reusability and sharing is better than with the database approach, thanks to the platform independence of XML. However, XML as a meta-language defines only the general syntax without formally defined semantics, which leads to difficulties when reasoning.

As an alternative to the above approaches for user modeling, the ontology-based approach has emerged. The benefits of using ontologies for user modeling and adaptation have been recognized by many researchers nowadays. The advantages leading to the use of ontologies for user modeling come from the fundamentals of this formalism. Ontologies provide a common understanding of the domain to facilitate the reuse and harmonization of different terminologies [17]. They support reasoning, which is considered an important contribution of the ontology-based models. Once user characteristics are in ontological representation, the ontology and its relations, conditions, and restrictions provide the basis for inferring additional user characteristics. For example, considering a user who is a programmer and works for a company that develops web-based applications using Java technologies, we can infer that she is skilled in Java technologies.

3. Adaptive System Supporting Collaborative Learning Based on Location-Based Social Network and Semantic User Modeling

3.1. Construction of Social Network Education Based on Location Information. E-learning is a more flexible and innovative way of learning. Recently, e-learning has become more popular. However, it is difficult to provide interaction services in real time. The existing e-learning methods only provide bulletins for sharing lectures or addressing the needs of study groups. They do not provide various methods for forming a community or a small group among the learners. In particular, it is inefficient compared with face-to-face offline meetings and discussions, and, thus, there exists limitations on collaborative learning.

For this purpose, in this paper, we will collect location-based information through the GPS sensor of a smart-phone. That is, it will support a method that can construct a social education network based on the collected location information of the learners. This method will inform a mobile learner of the current location or location logging information of other learners who are attending the same online course. Based on the location information provided to mobile learners, collaborative learning can be requested of other learners. If other learners in the vicinity are also attending the course, a learner may request to meet them at a specific place for discussion and collaborative learning. Figure 1 explains the basic concepts of this system.

In Figure 1, Learner A executes the online course application on a smart-phone. Learner A requests learning content from an e-learning server. The e-learning server generates additional information from the location information of Learner A and provides it together with the learning content.

The location of Learner A is tracked by the GPS sensing capability of the smart phone. The location information of Learner A is used to help him meet, at a specific place, other people who attend the same course if they are nearby. In this way, e-learning can be extended to face-to-face meetings. In other words, close social relationships between learners can be formed by using smart phones. Learner A sends a message requesting collaboration to the mobile-learning server for a person in the local area who can assist in the course. The

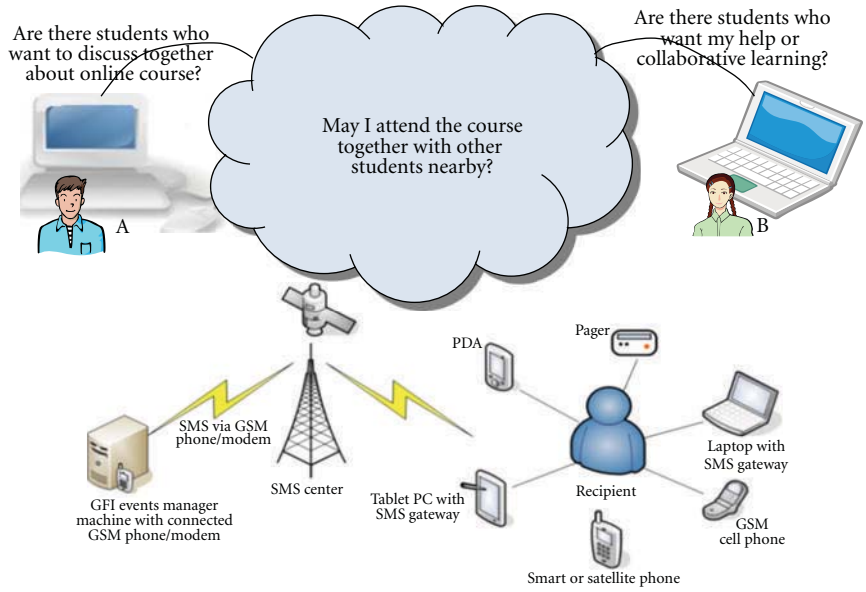


Figure 1: Construction of location-based social education network.

mobile-learning server sends the request message it received from Learner A to Learner B. It realizes a real-time social network by receiving a response to the request message.

However, for learning purposes, simply using the location information may not result in an optimal solution for the learning perspective. Therefore, the algorithm for constructing a social network requires further learning factors such as the learner's learning profile, learning style, and learning interests. The learner's learning style is considered an important constructing criterion. Learners with similar learning styles tend to have more interactions with one another during their learning experience. The constructing algorithm also considers a learner's learning interest as the learning criteria in addition to the learner's learning profile. These learning factors are related to user modeling information.

3.2. Semantic User Modeling. User modeling should be considered so as to form a social network for collaborative learning and to support adaptive learning. This study uses onotologies for user modeling that represents the characteristics of each learner. As mentioned in the previous section, the main reason for using ontology-based modeling is because it supports reasoning. The feature of reasoning is useful in providing personalized learning and collaborative learning depending on the learners' characteristics. Figure 2 shows the User Model ontology in the system provided in this study. It includes each learner's learning profile such as learning history, learning level, prerequisite knowledge, learning goals, learning time, and learning context, learning interest and learning style as shown in Figure 2. Learning context includes the operating system, terminal type, and network information for the learner's computer. Learning style is based on the four dimensions suggested by Kay and Lum [17]. This ontology-based user modeling supports semantic searches as the ontology and its relations, conditions, and restrictions in ontological representation provide the basis for inferring additional user characteristics.

3.3. System Framework. This system supports both personalized learning and collaborative learning. Personalized learning provides adaptively learning content according to each learner's level of knowledge, preferences, and other considerations. Personalization is essential for improved learning experiences. The adaptive learning system we propose provides highly individualized learning content based on ontologies. In addition, the system provides collaborative learning among users. However, it is important for learners to be properly grouped for collaborative learning. That is, learners would need teammates who share the same interests, preferences, or learning experiences. Ontology-based user modeling is useful for grouping learners together in collaborative learning as well as in personalized learning. Finally, this system supports face-to-face collaborative learning based on location-based sensing information.

The architecture of the proposed ontology-based adaptive learning system consists of three main ontologies (Domain, Content Structure, and User Model ontology) and four modules (Learning Content Management, User Model Management, Adaptive Learning Support, Collaborative Learning Support, Location-based Information Management module) as shown in Figure 3.

3.3.1. Learning Content Management Module. The learning content management module manages the repository of learning objects. It inserts a new learning object into the repository or manipulates existing learning objects. Each

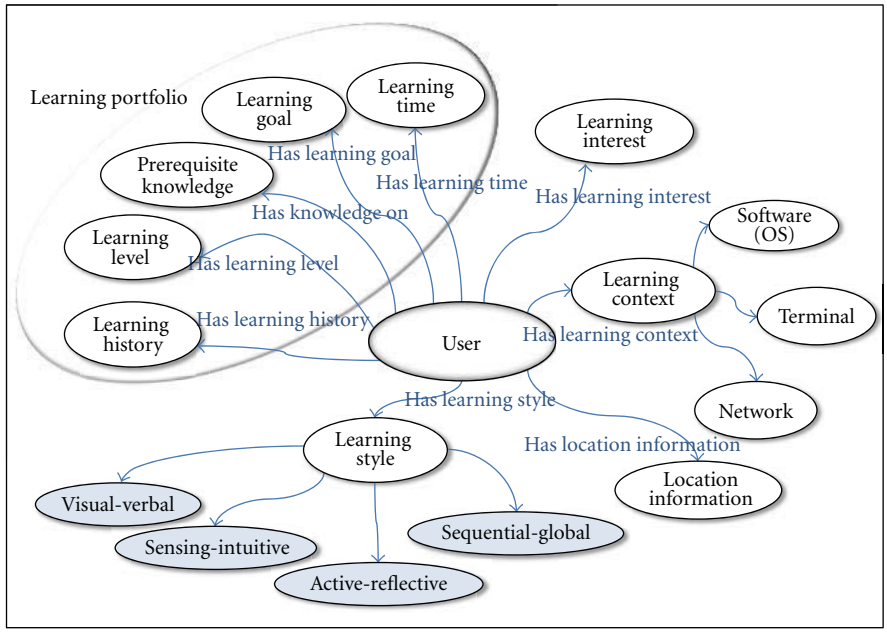


Figure 2: User model ontology.

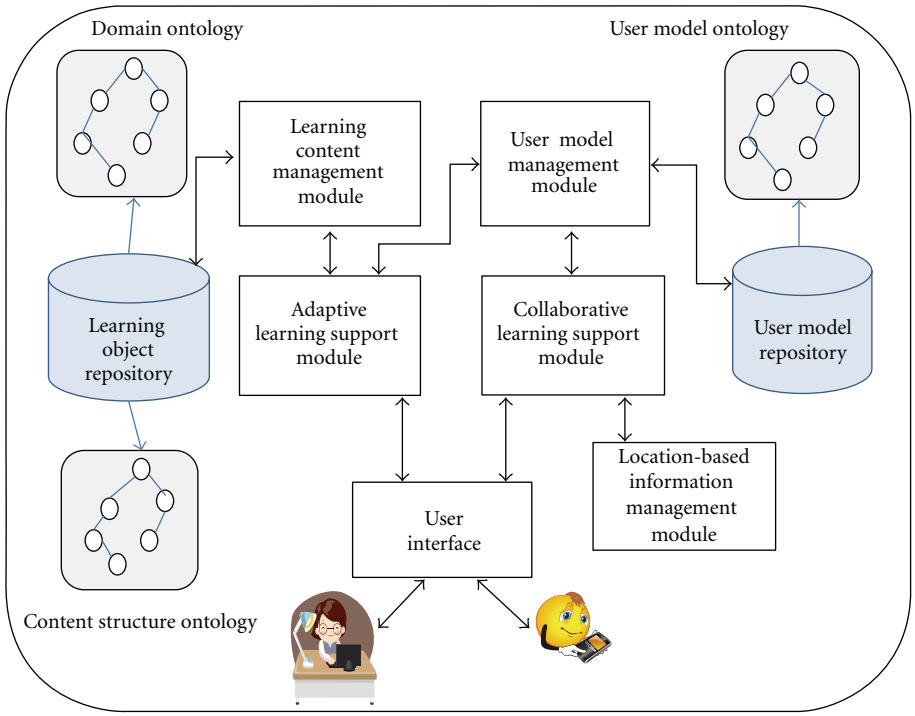


Figure 3: System architecture.

constructed learning object corresponds to each learning concept in the hierarchy of the content structure ontology. A parent object generally consists of a number of subchild concepts.

3.3.2. User Model Management Module. This module manages the repository of user models, responding to user information requests such as inserting into, updating, or accessing the repository.

The learner model is required to provide learners with useful information and can be constructed from learners' learning portfolios. Therefore, the system needs a learner model to track and represent from the learner's learning status. The learner model contains the learning status of every

learning concept and the learning sequences for their goals. Consequently, the learner model can support personalized learning for a specific learner and collaborative learning for peer-consultation. The learner model mainly consists of the learning status of each concept, the learner's learning goal, and learning time. In capturing the learners' learning status for each concept, all clues are generated by analyzing the learners' online learning behaviors and testing the results. The learner's model is built with them.

3.3.3. Adaptive Learning Support Module. This module dynamically generates personalized learning content depending on each learner. In association with the User Model Management module, it diagnoses the knowledge level of a learner based on a questionnaire and learner test results. It also analyzes the learner's learning history, learning context, and learning style. Based on these acquired learner characteristics, it guides the learner's learning process using the most appropriate learning content. We have also defined some inference rules to support adaptive learning.

This module is responsible for presenting personalized learning content according to each learner's characteristics. The system presents adaptive learning content using link-hiding techniques. If it analyses a learner's learning level and considers certain learning content as inappropriate for a learner, the link to that learning content is hidden from the learner. For example, if a learner has not learned about or does not have a good understanding of a concept such as the "Fertilization of Animals," the links to "Internal Fertilization" and "External Fertilization" are hidden so that the learner cannot select those concepts.

3.3.4. Collaborative Learning Support Module. This module supports constructing a social network for collaboration according to the learners' information in User Model ontology. The constructing algorithm creates the learning group for the learners with similar learning profiles and learning interests to facilitate their learning in pursuing their learning objectives. This constructing algorithm refers to an online learning group. Sometimes, learners will want to have face-to-face collaborative learning offline. For this to occur, this module provides constructing location-based social networks with the support of Location-based Information Management module. The constructing algorithm for location-based social networks is leaner centric in order to protect online mobile learners' privacy. A learner has to initialize the grouping request for collaborative learning. Otherwise, the learner's location and other data will not be used by the algorithm and shared with other learners. The learner who initialized the grouping request for collaborative learning can request face-to-face collaborative learning and contact other learners who are nearby.

3.3.5. Location-Based Information Management Module. The Location-based Information Management module receives the location information of the learners' smart-phone. The Location-based Information Management module collects the location information of the smart phone when there is a request for learning content or collaborative learning from a user. It can generate location-based learning information based on the collected location information or location log. Location-based learning information includes the information concerning the area and location where the learning content is being used. For example, it includes the information concerning how many times the learning content is used in a specific place such as a café, restaurant, library, park, and classroom. In particular, it may include the information regarding the specific address of each location (latitude and longitude by GPS as well as information about the administrative district). If the learning content is used frequently in one location, it means that the location is likely to become an appropriate place in which to use the learning content. Therefore, through additional information based on location, the learner can use the learning content more conveniently by finding a place where the learning content is frequently used. Alternatively, the learner can attend the course at a specific place or check whether it is an appropriate place to study or take a quiz.

The system shares the location information of the smart phone of learners who are attending the same course and want to have collaborative learning offline. In this way, the location information of the learners' smart phone located within the preset distance from the learner's smart-phone who requests collaborative learning is searched for and the location-based additional information, including the searched location information, is generated. When there is a request for collaborative learning from a learner, the Location-based Information Management module searches other learners who are near the learner. When a response approving the collaboration request is received, the detailed location information of the requested user's smart phone is sent together with the response to the learner's smart-phone. Algorithm 1 shows the algorithm that executes this process.

The learner's smart-phone application provides real-time service of the received location-based learning information to the learner. The service is made in a social display form combining AR (Augmented Reality) and the location information. The learner selects, among the displayed locations of other learners' smart-phones, those learners with whom he wants to attend the course and those who can provide assistance in the course. To the selected people, a message requesting a learning collaboration, and so forth, is inputted to the learner's smart-phone. The learner of the smart-phone which received the collaboration request may send a response accepting the collaboration request to the system. The system sends the response to the collaboration request to the learner's smart-phone. If the response to the learning collaboration request is acceptance of the request, the system sends the detailed location information of the accepting smart-phone to the smart-phone of the learner. The locations of the learner can be easily recognizable on Google Maps (http://maps.google.co.kr/), Naver Map (http://map.naver.com/), or Daum Map (http://local.daum.net/). Therefore, a face-to-face meeting between an instructor and a learner who are nearby is possible. Furthermore, the additional information on the frequency of a learning collaboration or the frequency of learners visiting

```
/* c_l_i DB: Course Learner/Instructor DB */
   Get a Request_id
   If (the Request_ID == Terminal_ID in course Learner/Instructor DB) then
        Get the Terminal_location and Course_ID of Request_ID
   EndIf
   For each record in c_l_i DB Where (record.Course_ID == Request_ID.Course_ID)
      Compute the Relative_location of the Terminal_ID
      If (the Relative_location <= θ) then
           Get the Terminal_IDs
      EndIf
   EndFor
   If {Terminal_IDs} ≠ null then
      For each Terminal_IDs where (Relative_location <= θ)
         Send a request for social network to each Terminal_ID
      EndFor
   EndIf
   If    ∃(acceptance signals for social network from the Terminal_IDs) then
      Send signals with detail location information to the Request_ID
   EndIf
```

Algorithm 1: Location-based information management module's execution algorithm.

Figure 4: The search display of learners and instructors with whom learning collaboration is possible.

a specific place (café, restaurant, library, park, bookstore, etc.) is shared. Learning collaboration helps form social relationships through twitter (http://twitter.com/), facebook (http://www.facebook.com/), me2day (http://me2day.net/), yozm (http://yozm.daum.net/), and so forth.

4. Service Scenario

Smart-phones can obtain a more-detailed context awareness in the mobile environment through sensors such as GPS, Bluetooth, Ambient light sensor, Approximate Sensor, Accelerometer (gravity-sensor, Gyro sensor, CandleFrame, BobbleMonk, Punch Mach, Neon Board), Digital Compass (Magnetic Field Sensor), and WLAN.

The location information obtained by sensing is used to recommend the learners or the instructors located nearby the current location of the learner or to provide a social network service. Figure 4 displays the learners and the instructors located nearby on Naver Map (http://map.naver.com/). That is, Figure 4 is the display that searched for among the learners, other learners located within "100 m," "50 m," "1 km,", and so forth from the location of the current learner. The

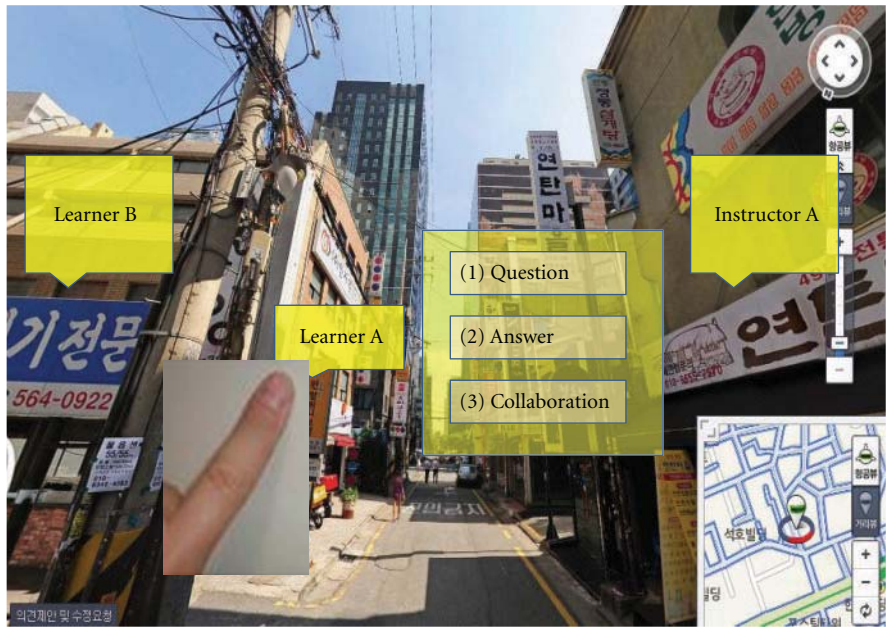


Figure 5: Augmented reality based on social learning and location information.

learner can suggest a learning collaboration to the nearby learners or the instructors displayed on the smart-phone. This makes it possible for the nearby learners and instructors to meet face-to-face for learning.

Figure 5 is the screen displaying the learner and the instructor on Naver Map (http://map.naver.com/) through Augmented Reality technology. A learner with a smart-phone can realize Augmented Reality through a camera module. By combing Naver Map and Augmented Reality technology, the service helps make it easy to find the learners or the instructors located nearby. Therefore, a social relationship based on location becomes possible. At the same time, it can be extended to social network services such as Facebook, Twitter, Me2day, and Yozm.

For learning in the mobile environment, location information can assist the learner in recommending the best colleagues or selecting the best place for learning from the actual location of the learner. Moreover, it helps in obtaining learning-related information on the Internet through social network services. A mobile learning process such as this naturally leads to meeting other friends face-to-face and solves the problem of isolation in online learning activities in virtual reality. In particular, location-based learning information helps make the learner's personalized learning possible in a context awareness environment.

5. Conclusions

Distribution of smart-phones provides a mobile learning environment. This paper proposes a method for combining the location-based information of smart-phones and e-learning, and realizing an e-learning support system based on location-based social network. The proposed system supports collaborative learning based on a location-based social network and semantic modeling. That is, a social network among e-learning learners is dynamically constructed on the basis of the location information of the learner using GPS sensors thus enabling face-to-face collaborative learning. In addition, the system offers adaptive help and advice regarding the most effective and productive organization of learners in groups for collaborative learning. This is done so that the users' collaboration can be more efficient. Through this system, e-learners can find mentors who facilitate their learning and create communities for collaborative learning off-line and on-line. The support of a location-based social network service and semantic user model in our system would increase interactions among e-learners and improve satisfaction regarding their mobile learning environment.

Acknowledgment

This work was supported by Woosuk University (2012).

References

[1] E. Fredericksen, A. Pickett, P. Shea, W. Pelz, and K. Swan, "Learner satisfaction and perceived learning with on-line courses: principles and examples from the SUNY learning network," *Journal of Asynchronous Learning Network*, vol. 4, no. 2, pp. 7–41, 2000.

[2] L. Y. Muilenburg and Z. L. Berge, "Learner barriers to online learning: a factor analytic study," *Distance Education*, vol. 26, no. 1, pp. 29–48, 2005.

[3] V. Tinto, "Learning better together: the impact of learning communities on learner success in higher education," *Journal of Institutional Research*, vol. 9, no. 1, pp. 48–53, 2000.

[4] S. Carr, "As distance education comes of age, the challenge is keeping learners," *Chronicle of Higher Education*, vol. 46, no. 23, pp. 39–41, 2000.

[5] D. Boyd, "The characteristics of successful on-line learners," *New Horizons in Adult Education*, vol. 18, no. 2, pp. 31–39, 2004.

[6] I. Liccard, "The role of social networks in learners' learning experience," *SIGCSE Bulletin*, pp. 224–237, 2010.

[7] Q. Tan, K. Yu-Lin, and Y. M. Huang, "A collaborative mobile virtual campus system based on location-based dynamic grouping," in *Proceedings of the 10th IEEE International Conference on Advanced Learning Technologies (ICALT '10)*, pp. 16–18, Sousse, Tunisia, July 2010.

[8] Y. Zheng, Y. Chen, X. Xie, and W. Y. Ma, "GeoLife2.0: a location-based social networking service," in *Proceedings of the 10th International Conference on Mobile Data Management: Systems, Services and Middleware (MDM '09)*, pp. 357–358, Taipei, Taiwan, May 2009.

[9] M. Li and Z. Du, "Dynamic social networking supporting location-based services," in *In Proceedings of the International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC '09)*, pp. 149–152, Hangzhou, China, August 2009.

[10] C. M. Chen and Y. N. Tsai, "Interactive location-based game for supporting effective English learning," in *Proceedings of the International Conference on Environmental Science and Information Application Technology (ESIAT '09)*, pp. 523–526, Wuhan, China, July 2009.

[11] P. Kinnunen and L. Malmi, "Problems in problem-based learning—experiences, analysis and lessons learned on an introductory programming course," *Informatics in Education*, vol. 4, no. 2, pp. 193–215, 2005.

[12] R. Wegerif, "The social dimension of asynchronous learning networks," *Journal of ALN*, vol. 2, no. 1, pp. 34–49, 1998.

[13] D. D. Curtis and M. J. Lawson, "Exploring collaborative online learning," *Journal of Asynchronous Learning Network*, vol. 5, no. 1, pp. 21–34, 2001.

[14] S. R. Hiltz, N. Coppola, N. Rotter, and M. Turoff, "Measuring the importance of collaborative learning for the effectiveness of ALN: a multi-measure, multi-method approach," *Journal of Asynchronous Learning Network*, vol. 4, no. 2, pp. 103–125, 2000.

[15] C. Liao, F. Chen, and T. Chen, "Perspectives of university learners on cooperative learning by moodle," *International Journal of Digital Content Technology and Its Applications*, vol. 5, no. 6, pp. 190–197, 2011.

[16] A. Andrejko, M. Barla, and M. Bielikova, "Ontology-based user modeling for web-based information systems," *Information Systems Journal*, vol. 2, pp. 457–468, 2007.

[17] J. Kay and A. Lum, "Ontology-based user modeling for the semantic web," in *Proceedings of the 10th International Conference on User Modeling (UM '05)*, pp. 11–19, 2005.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 247346, 11 pages
doi:10.1155/2012/247346

Research Article

CA_{5W1H}Onto: Ontological Context-Aware Model Based on 5W1H

Jeong-Dong Kim, Jiseong Son, and Doo-Kwon Baik

Department of Computer and Radio Communications Engineering, Korea University, Seoul 136-701, Republic of Korea

Correspondence should be addressed to Doo-Kwon Baik, baikdk@korea.ac.kr

Received 11 October 2011; Accepted 24 December 2011

Academic Editor: Tai Hoon Kim

Ubiquitous computing is necessary to define models for broad contextual information arising out of surrounding environment. It also helps comprehend how to model a mechanism of selectively collecting useful pieces of contextual information and of providing relevant intelligent services. Further, studies are also required on how to process contextual information, its maintenance, and reasoning. However, current context-aware researches are still in need of modeling techniques reflecting ontological characteristics. As a result, it is impossible to effectively provide relevant intelligent services. They are limited as well in terms of contextual reasoning and interoperability across different pieces of contextual information. Aware of the issues, this study proposes an ontology-based context-aware modeling technique, along with a relevant framework, in order to enable efficient specification of contextual information and, thereby, further to provide intelligent context-aware services for context management and reasoning. Moreover, we mobilize the maxim of "five Ws and one H" to process physical and logical contextual information and to support our proposed technique. The maxim-applied modeling technique sets forth an intuitive context-aware schema and demonstrates high applicability to sharing and integration of contextual information. Meanwhile, the ontology-based modeling supports reasoning on contextual information and facilitates more intelligent and reliable services.

1. Introduction

Ubiquitous computing is an environment unbounded by spatiotemporal constraints. It enables computing among humans, objects, and information through interaction among diverse devices immersed in the surroundings. To realize it, support should be secured for context-aware technology pertaining to environment, time, space, and user. One of the gravest problems to such realization lies in lack of the technology that manages and processes various situations arising out of the ever-changing ubiquitous environment. Context-aware technology roots its core in the ability to conduct self-adjustment of conditions, or the ability to detect changes to conditions and to deliver relevant information to users or their handsets [1]. The notion of "context" is dividable in two subcategories: explicit context and implicit context. The former refers to the information that clearly shows the properties of the context, while the latter indicates the collected information that does not reveal its characteristics. The components used for distinguishing the two are concerned not only with the properties of contexts themselves, but also with the devices detecting the contexts. Location-tracking data, for instance, become an explicit context only upon widespread use of global positioning system (GPS) devices.

Even in this case, the data locating a user at a hospital constitutes an explicit context, while the purpose of her visit remains implicit. The purpose, however, turns explicit upon availability of access to hospital reservation system and use thereof.

It is vital to realization of this context-aware technology to be able to provide such environment and technologies that invisibly and automatically process the contextual information received from sensors. In other words, it is necessary to design an intelligent context-aware model for computers' self-detecting and -processing of contexts.

Under conventional ubiquitous computing circumstances, application of context-aware technology requires intervention of agents in order to provide services best tailored for each situation [2–6]. Aware of the barrier, studies have continued to find breakthroughs. For example, researchers are currently looking into the topic

of context-aware reasoning-supporting mechanism under which ontology-based devices carry out, for themselves, interoperability of different contexts and semantic analysis, as the context-aware technology advances based on intelligent sensors, devices, and systems [7]. Application of ontology to current ubiquitous computing environment is still faced with the issues shown below, despite the numerous studies on how to ontologically process context-aware information.

(i) First, how to form a context-aware model commands the utmost importance in developing an ontology-based framework. The strong coupling context-aware model excels due to its domain-specific quality services, while it requires additional mechanisms to process context-aware information in other domains. Likewise, a loosely coupled context-aware model is beneficial for its ability to use common contexts in processing various types of contextual information. It is impossible under loose coupling, however, to induce reasoning as to how to perform context application and combination to provide quality services.

(ii) Another problem is how to maintain interoperability over context-based data produced by diverse devices and sensors.

(iii) Finally, it is still a remote wish to provide perfect support for functions such as context reasoning for context combination. Especially, solutions must be found as to context application (e.g., services for the context B produced via the context A) and context combination (e.g., the context D as sum of the contexts B and C under the context A) in order to support intelligent, tailored services under the ubiquitous computing circumstances.

To address the problems, it is necessary to study how to define an efficient model that calibrates abstractness of the broad contextual information existing in ubiquitous computing environment. At the same time, it is also worthwhile to study how to model a mechanism of selectively collecting useful pieces of contextual information and providing relevant intelligent services. Further, studies are also required on how to process contextual information, its maintenance, and reasoning.

Thus, in this paper, we define an ontology-based modeling technique upon application of the "five Ws and one H (5W1H)" maxim for clear definition and management of contextual information and further propose a framework that supports context management and reasoning. The framework proposed herein defines a maxim-applied ontology-based modeling technique (i.e., context-aware model based on ontology using five Ws and one H: $CA_{5W1H}Onto$) as a method for interpretation and abstraction of semantic contexts. It is possible to define the $CA_{5W1H}Onto$ by applying formalism, which is used as a notion of ontology in connection with expressiveness of contextual information. That way, complexity is easily definable by means of ontological properties. In other words, the $CA_{5W1H}Onto$ provides some formalism as to contextual information via the ontology OWL-DL (web ontology language-description logic) [7]. In addition, with the use of the ontology OWL-DL, it is possible to model a specific domain by defining classes, instances, datatype properties (characteristics of instances), and object properties (relations between instances). Further, complexity descriptions of classes and properties are set forth by composing elementary descriptions through specific operations provided by the OWL.

The rest of this study is organized as follows: Section 2 describes the related work, and the concept and framework of the $CA_{5W1H}Onto$ is illustrated in Section 3, Section 4 describes its experiment and implementation, and Section 5 concludes this paper with recommendations for future work.

2. Related Work

Latest developments in the field of the semantic web have unfolded a new set of applications. The semantic web has enabled use of the tools necessary to handle computer-understandable semantics. Generally evolving from XML, these tools function as vehicle to enrich the description of web pages and, thereby, help fully understand the relations between the concepts. OWL [7] and resource description framework (RDF) [8] represent some of the most widely used instruments. These languages offer a significant advantage. They are machine readable and strongly related to description logics. A state of the art on this subject is found in [9]. The RDF language produces resources-related statements in the form of the "subject, predicate, object" triple. The "subject" denotes the resource, and the "predicate" refers to the relationship between the "subject" and the "object." Built on top of the RDF, the OWL language offers a larger vocabulary and stronger syntax.

Numerous studies have been pursued to come up with modeling techniques for processing of contextual information. Especially, ontology-based context-aware models have the following merits in common: (i) they support ample relations through simple expressiveness; (ii) they enable sharing and integration of contexts belonging to different sources through formal semantics; (iii) they also enable consistent definition of relations to express contexts with ontological tools (e.g., protégé [10], TopBraid [11]); (iv) they further support automatic reasoning, in a consistent way, on expressiveness of various types of contextual information.

Hereunder, we will review some of the most relevant context modeling approaches. These are classified by the scheme of data structures, the structures that are used to exchange contextual information in individual systems.

Conventional key-value models [12, 13] take on the simplest data structure for context modeling purposes. A key-value model provides values of context information as environmental variables and conducts modeling by means of key-value pairing. It is easy to maintain the model. The model, however, is limited in its inapplicability to such structure as fit for efficient context-searching algorithms.

The markup scheme model is a hierarchical data structure consisting of markup tags retaining properties and contents [14]. Its markup tag is always reflexively defined by

other markup tags. The profile is a representative example of this technique. It operates off the serialization induced under standard generalized markup language (SGML), the "superclass" of all markup languages like XML.

The object oriented model is a technique that accommodates the major benefits of being object-oriented like capsulation and recycle in order to address the problems arising in connection with dynamic attributes of a context in the ubiquitous environment [3]. The details concerning the context processing are capsulated at the individual object level and hidden in other components, while access to context information is made only through a certain interface.

A logic-based model expresses, based on and through rules, the knowledge about various contexts and provides applications with the values of different contexts [15]. This logic-based approach is hard to maintain and consumes time and money to resolve conflicts, as the number of rules gets bigger.

The notion "ontology" as in the ontology-based models represents a vehicle used to describe concepts and correlations [16]. It provides vocabulary and terminology for sharing. Due to the diversity of context information, ontology is defined for each domain. This type of model supports sharing and recycle of context information in the ubiquitous computing environment. Meanwhile, the hierarchical ontology-based model creates lower-level ontology tailored for a certain domain, using the upper-level ontology. The ontology-based context-aware modeling technique, therefore, supports reasoning on context information based on correlations among context data. With the reasoning possible, a device conducts semantic analysis for itself.

In this and other related fields, several approaches have been proposed for designing architectures designed to mode users and minimize user interaction in the process of supporting situation-dependent personalization of services (e.g., [17]).

Traditional context-aware systems are to be introduced hereunder, to which ontology is applied. Some of the noticeable examples are context broker architecture (CoBrA) [16], service-oriented context-aware middleware (SOCAM) [18], middleware platform for active spaces (Gaia) [19], and context toolkit [20].

Supporting the ontological language "OWL [7]," CoBrA [16] is distinguishable from other systems in that it employs the intelligent space-based architecture. Space functions as agent for creation, management, and distribution of context information in the ubiquitous environment. CoBrA ontology was proposed for modeling context information through ontology. Designed in OWL, CoBaA-Ontology aims at developing an independent ontology for expression in a system. CoBrA-Ontology supports reasoning and context determination for transmission to web services. In detail, it defines ontology each for action, agent, device, conference, digital document, space, and time. Then, it uses the definitions to provide services tailored for tag and sensor values, or time and space of a user. It lacks interoperability with other instances of context-aware ontology. The shortcoming comes from its use of nonintuitive ontological properties in the ontology-based modeling, and from its focus on development of broker agent for sharing and maintenance of all contexts.

Based on service-oriented architecture (SOA), SOCAM [18] is a middleware running on open services gateway initiative framework (OSGi). Using OWL-based ontology, it supports reasoning on expressions about contexts and on various forms of them. Within the middleware, OWL encoded context ontology (CONON) [21] is used to transform context information into ontology. CONON defines as ontology the upper classes (i.e., ontology) retained by most context-aware systems for its easy application to multiple domains. The feature enables other domain systems to use the upper classes and easy sharing. Based on SOA, the objective of SOCAM is to resolve problems concerning interoperability, sharing, and classification of knowledge. It, however, employs a very complex mechanism for sensing of context information. Further, as loose-coupled ontological modeling technique, CONON carries limited functionality in terms of reasoning on context information (e.g., context application and combination) and high-quality services.

Gaia [19] is a type of the conventional operating system expanded for expressing context information. Gaia collects information on objects in use of ontology servers and collects context information through sensor networks. Then, it extracts upper context (i.e., higher-level context) via reasoning based on the collected information on objects and contexts. Based on DAML+OIL [22], the context model of Gaia takes on the form of "a subject and an object" with a verb inserted in between. Lack of definition and expression of verbs renders semantically perfect context-awareness impossible. Moreover, the DAML+OIL applied in Gaia is more limited than OWL in the inferring functionality for provision of intelligent tailored services.

Context toolkit [20] employs the object-oriented approach and enables easy use of contexts by providing recyclable modules and easily developable frameworks. Strictly speaking, context toolkit lies outside ontology technology. However, it uses the attribute-value tuples for modeling, which assume ontological nature. This mechanism is not an appropriate tool for the purpose of our study, because the absence of a reasoning engine poses limitation in providing context-aware intelligent services, and programming for processing of context information is too complex.

Indeed, most of the aforementioned studies propose ontology-based context-aware modeling techniques. Still, they focus just on defining diverse types of contextual information in ontological language. Further, they employ different modeling techniques for different systems and do not efficiently carry ontological properties (e.g., reasoning) concerning modeled context-aware ontology. Consequently, the studies have failed to overcome the limits pertaining to undefined contextual information. In addition, adoption of different modeling techniques has led to problematic expansion of interoperability and integration among systems.

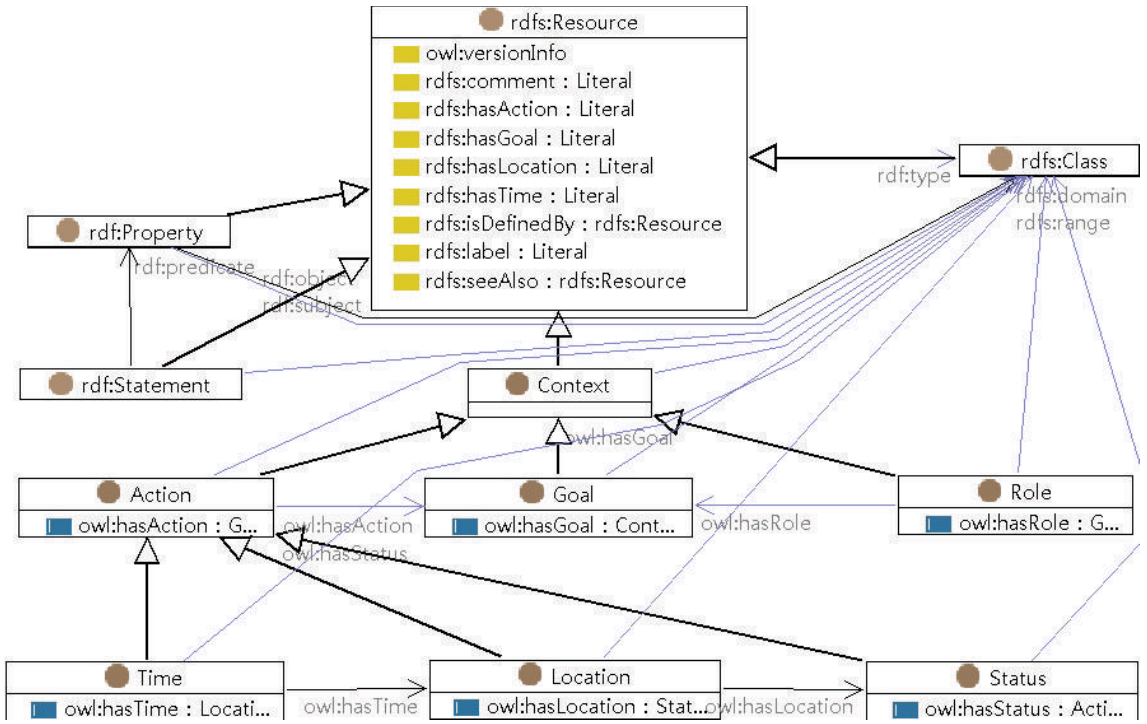


Figure 1: Expressing conditions of the context-aware.

3. Our Proposed Approach

Reference [23] defines context as *"Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves."* In other words, context refers to a special form of knowledge. It constitutes an important modeling requirement. One of the examples is the tradeoff between expressiveness and complexity.

The tradeoff is an important component in determining expressiveness of knowledge and reasoning capability. To resolve the tradeoff issue concerning the context-aware modeling, ontology (i.e., OWL-DL) may be utilized. Ontology amply expresses concepts and their relationships and automatic reasoning in processing contexts based on the expressive capacity.

Proposed herein is a context-aware ontological model based on five Ws and one H (i.e., CA_{5W1H}Onto) as a method of interpreting and abstracting semantic contexts. Our CA_{5W1H}Onto is designed to support intuitive integration of different context-aware schemas, to which the maxim is applied.

The proposed model consists of <`Concept`, `Instance`, `Context`> triples, and the first two elements of the triple set utilize the properties defined through the existing ontology. The `Context` element carries all six attributes of the maxim, or why, who, what, where, when, and how. The elements are, in turn, mapped onto corresponding ingredients of the context-aware processing (i.e., goal, role, action, status, location, and time). Figure 1 shows the key elements constituting our proposed model and their relations in class diagram.

3.1. Context-Aware Framework. Figure 2 illustrates the proposed context-aware framework consisting of two parts: "Context Modeling based on ontological concepts" and "context managing and reasoning."

The "context modeling based on ontological concepts" (upper part in Figure 2) is named CA_{5W1H}Onto. The CA_{5W1H}Onto model consists of `Concept`, `Instance`, and `Context`. The element of `Context` contains various contextual activities. In turn, the `Context` component helps define the basic characteristics of the maxim and the schemas mapped among context-aware elements. The CA_{5W1H}Onto model proffers services tailored for a specific time, space, and set of user preferences across different domains. The model performs modeling of essential elements by defining, in accordance with the maxim, the contexts required for integration and interoperability of the defined contextual information.

The CA_{5W1H}Onto model defines, in the unit of <`Concept`, `Instance`, `Context`>, the ontological elements (e.g., concept, instance, datatype, data property, object property, etc.) and the elements (i.e., goal, role, action, status, location, and time) for context-aware definition. By defining each in an independent component module, adaptability and independence are guaranteed when developing a context-aware model applicable to diverse domains. In other words,

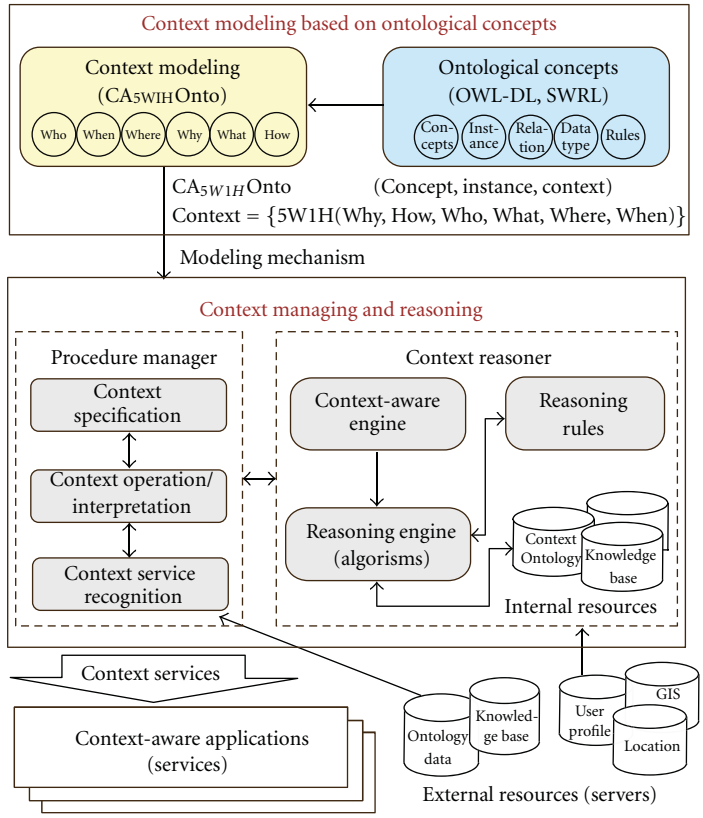


Figure 2: Proposed context-aware framework.

the maxim-applied context-aware modeling technique is an intuitive model in nature and thereby enables interoperability between systems or models throughout integration and sharing of schemas belonging to diverse domains. Detailed explanation ensues hereunder as to how to map between the properties of the maxim and the context-aware elements defined by the CA5W1HOnto model.

One of the merits from ontologically modeling contextual information lies in its capabilities of automatically extracting new knowledge about current contexts, and providing ample formalism with streamlined expressiveness about the knowledge. Incorporating the merits in the CA5W1HOnto model for support of better intelligent tailored services requires defining algorithms and rules for context reasoning and managing procedures.

"Procedure manager" and the "context reasoner" of the proposed context-aware framework are the functions taking charge of "context management and reasoning" (lower part in Figure 2), respectively. The former manages specifications, processes, and services concerning the contexts defined through the CA5W1HOnto model, consisting of "context specification," "context operation/interpretation," and "context service recognition." In the meanwhile, the latter consists of "context-aware engine," "reasoning algorisms," and "reasoning rules," the components that support context application and reasoning for provision of better intelligent context-aware services.

In this light, the proposed context-aware framework enables combination of high-level context-aware services via efficient management and application of contextual information. Due to the scope of this study, we focus on context-aware modeling and process, leaving for the future such topics as reasoning algorithms and rules for context management and reasoning.

3.2. Modeling of Context Using 5W1H. To process contextual information used across distinct domains requires a model designed to process very complex and multiple contextual situations. Developing this type of model, utilization of the maxim of five Ws and one H helps share and integrate contextual information and, eventually, helps provide quality context-aware services. This section explores, in detail, the proposed CA5W1HOnto model as a context-aware, maxim-applied modeling technique.

Table 1 demonstrates how the CA5W1HOnto constituting maxim attributes are mapped onto the ingredients required for context-aware definition. Specially, the CA5W1HOnto model retains, to define contexts, the six elements of `role`, `goal`, `Action`, `Status`, `location`, and `time`. In addition, the mapping is coupled as `why::Goal`, `who::Role`, `how::Action`, `what::Status`, `where::Location`, and `when::Time`. Under the proposed context-aware modeling, `why::Goal` represents a desired state that one may wish

Table 1: Mapping between Context-Aware Elements and Maxim Attributes.

Construct in maxim-applied elements	Construct in context-aware elements	Definition
Why	Goal	A reason or goal that explains why a context processing has occurred
Who	Role	An entity that processes roles, namely, an agent that may be a person, an organization or a system involved in a context
How	Action	An action leading to the contexts, namely, a context may occur, when it is acted upon by another entity that is often a human, or a software agent
What	Status	A context (i.e., change of state) that happens to data during its lifespan (gathering of contextual information via various events)
Where	Location	A location associated with a context
When	Time	Time, or more accurately the duration of a context

to pursue, while its `who::Role` and `how::Action` indicate actual steps to be taken to achieve the state.

(i) `why::Goal`: as the ultimate goal of our context-aware processing, the coupling of `why::Goal` corresponds to the "why" property of the maxim, and is divided into *Personal_Goal*, *Functional_Goal*, *NonFunctional_Goal*, and *Role_Goal*.

(ii) `who::Role`: it defines the role necessary for processing of `who::Role`, stipulating the roles of *Role_Goal*, *Organization*, *Actor*, and *System*. It represents the "who" property. It further retains *Profile* and *SW_Agen* as subclasses required for defining the role of actor.

(iii) `how::Action`: corresponding to the "how" attribute of the maxim, `how::Action` shows how context-aware services are processed. In other words, it represents the context-aware processing stage leading up to realization of `why::Goal`. Its ingredients are *Atomic_Action*, *Composite_Action*, *Expectation*, *Precondition*, *Effect*, *Input, Output*, and the like.

(iv) `what::Status`: it represents unique attributes of a context object, which have been detected through sensing. This coupling corresponds to the "what" property and is expressed with *Atomic_Status* and *Composite_Status* ingredients.

(v) `where::Location`: The `where::Location` corresponds to the "where" property and provides location-related information.

(vi) `when::Time`: the `when::Time` represents the "when" attribute and uses *Start_Time*, *End_Time*, and *Repetition_Time* for context-aware processing.

The overall structure of the CA_{5W1H}Onto modeling is presented in Figure 3. In addition, Figure 3 illustrates how contextual information is processed, upon application of our proposed CA_{5W1H}Onto modeling technique.

Processing of contextual information starts with determination of which properties unique to an object are likely to be sensed (i.e., selection of sensed data). This phase corresponds to the "what" property of the maxim and is processed at the `Status`. Once selected, the `Location` and the `Time`, which represent the "where" and the "when" attributes, respectively, provide the time and location information on the data sensed at the `Status`. In charge of how to process the sensed contextual information, the `Action` remits to the `Goal` such property-processing results as *precondition*, *expectation*, *effect*, *input*, and *output*. In this case, the *Precondition* class preliminarily processes the constraint(s). Finally, the `Goal` sets forth the services necessary to process and define contexts, based on the processing results obtained at the `Action` and the `Role`.

Set forth hereunder for our proposed CA_{5W1H}Onto model are the definitions of the ingredients, the maxim attributes and their mapping relations, and the context-aware modeling techniques.

Definition. CA_{5W1H}*Onto Model*

$$\begin{aligned}
&\textbf{CA}_{\textbf{5W1H}}\textbf{Onto} = \{Concepts, Instances, Contexts\},\\
&\textbf{Context Model} = \{\textbf{5W1H}(Who, Why, How, What, Where, When,)\},\\
&Who = \{\textbf{Role}(Actor(Profile), Organization, System)\},\\
&Why = \{\textbf{Goal}(Personal_Goal, Role_Goal, Nonfunctional_Goal, Functional_Goal)\},\\
&How = \{\textbf{Action}(Atomic_Action, Composite_Action, Expectation(Failure), Input, Output, Effect, Precondition, contextual, trustworthy)\},\\
&What = \{\textbf{Status}(Atomic_Status, Composite_Status)\},\\
&Where = \{\textbf{Location}(Atomic_Location, Composite_Location)\},\\
&When = \{\textbf{Time}(Start_Time, End_Time, Repetition_Time)\}.
\end{aligned}$$

The contextual information that has undergone modeling through the six steps goes on through another three processes for context management at the "context managing and reasoning," as shown in Figure 2: "context specification," "context operation," and "context service recognition." Going

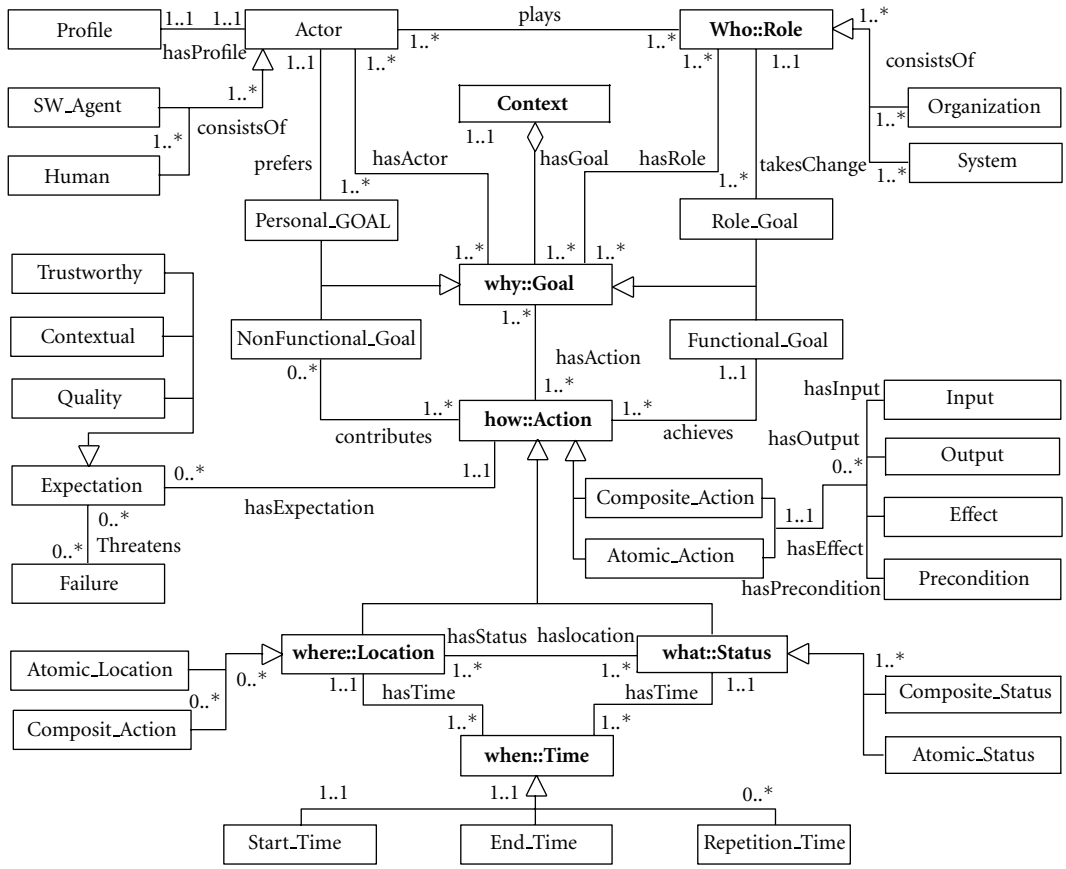


Figure 3: CA_{5W1H}Onto (ontological context-aware using 5W1H) model.

through the three processes of processing and selection, the contextual information leads to provision of significant context-aware services. Such contextual information-processing technique, upon application of the five Ws and one H maxim, provides a more intuitive model as to the information and serves as criterion for integrating context-aware data used in diverse domains. Thereby, it enables consistent integration and application of the information in defining and managing it.

4. Experiment

The CA_{5W1H}Onto has been defined in OWL, the web ontology language. As ontology editor, we have used the TopBraid composer [24]. In order to ensure our approach intuitive and understanding of the ontology simple, we have assumed the simple scenario as described in Section 4.2.

The TopBraid Composer is an enterprise-class modeling environment for developing semantic web ontologies and building semantic applications. Fully complying with W3C standards [24], composer offers comprehensive support for developing, managing, and testing configurations of knowledge models and their instance knowledge bases. As part of TopBraid suite, composer incorporates a flexible and extendible framework with a published API to develop semantic client/server or browser-based solutions that can integrate disparate applications and data sources.

Furthermore, in order to enable broad tool support, and to ensure computational completeness and decidability, the ontology has been developed with the aim to conform to the OWL-DL subset of the OWL ontology language.

4.1. Classes and Properties. Figure 4 illustrates the class hierarchy of the ontology, whereas Figure 5 shows the list of created properties, object properties (blue), datatype properties (green), annotation properties (yellow) for the CA_{5W1H}Onto. The class hierarchy splits the classes into two groups, or "context" classes and "ontology" classes.

The category of the "context" contains definitions of 39 classes like goal, action, status, and location. Thus defined, the classes are fixed, since they are vital to context-aware processing.

The ontology class sets forth definitions of *Car_Server*, *Conference*, and *Emergency_Center*. The proposed CA_{5W1H}Onto model defines ontologies and contexts after separating them in the form of module. Therefore, the predefined ontologies become adaptable to the model without modifications. In other words, additional ontologies

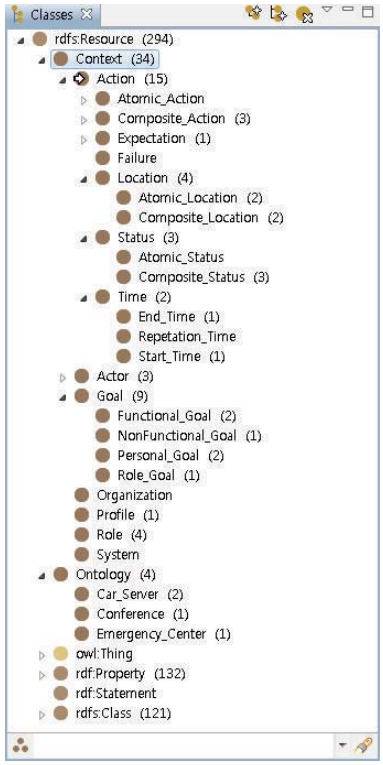


Figure 4: Class hierarchy in CA_{5W1H}Onto.

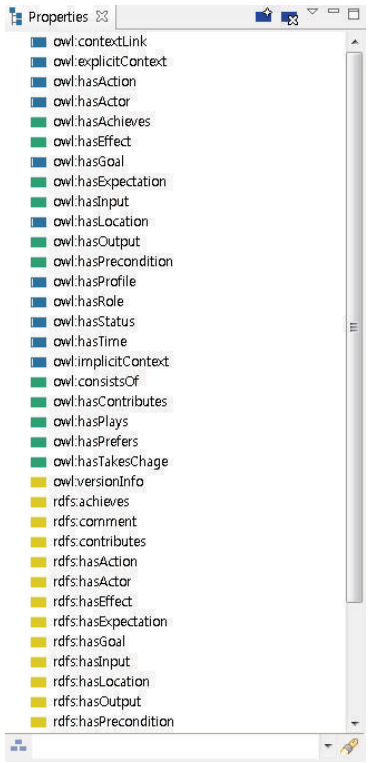


Figure 5: List of properties in CA_{5W1H}Onto.

are flexibly addible to each domain for context-aware processing.

4.2. CA_{5W1H}Onto Definition. In this chapter, a usage example is offered, based on a simplified scenario, to attest to the validity of our proposed CA_{5W1H}Onto model. Specifically, it is shown how our modeling technique works and how contextual information is processed.

A circle in Figure 6 represents each class, and a rectangular indicates each property. Further, the rhombus indicates each instance. Hereunder, we will describe the created classes, object, and datatype properties and visualize their relations in TopBraid-OWL editor's graph.

Scenario. Working for SYS. Lab, *Jane* hits the road to attend a *conference* scheduled at *3:00 pm* in city B. On the highway, however, she runs into an *autoaccident*. Sensing the impact through the *autosensors*, the *Car_Server* reads *GPS* coordinates with the *navigation* mounted on her vehicle and sends the information to the nearest *emergency center* and police station. At the same time, Jane's *scheduler* mounted on her handset makes inference on the situation and sends out her *absence notice* to the conference organizer.

The *Car_Server* (i.e., ontology) is related to the *Emergency_Center*, or another otology. In the former, various situations possible to occur to her vehicle are stored as ontology instances. When the relation between the two is processed as an explicit context "*Auto_accident,*" its impact affects the context that *Jane* previously pursues in connection with attending a conference. As a result, a new relation forms between the *Car_Server* and the *Conference* (i.e., ontology), and the new one is processed as an implicit context "*Inability_to_attend*."

The context arising between the *Car_Server* and the *Emergency_Center* is processed, based on the five Ws and one H maxim. `why::Goal` has *Personal_Goal* and *Functional_Goal* as its subclasses. Under the scenario, the explicit goal is to send out a notice on the autoaccident to the *Emergency_Center* by the *Car_Server*. At the same time, it becomes an implicit goal to notify the conference organizer of her failure to attend due to the accident, an act that she has been pursuing prior to the accident. Therefore, *Functional_Goal* is to notify the accident, while *Personal_Goal* is to inform the organizer of her absence. `why::Goal` forms a relation with `who::Role` as its object.

`who::Role` is as its subclasses actor and organization, with Jane as actor instance under the scenario. Meanwhile, `why::Goal` retains `how::Action` to understand what operations are in need among devices and `what::Status` that shows the device information. As subclasses of `how::Status`, *Atomic_Action* and *Composite_Action* exist with the former carrying out the process single handedly and the latter operating it in a composite way. Likewise, `what::Status` retains *Atomic_Status*, which single handedly processes subclasses, and *Composite_Status*, which carries it out in a composite way. The subclasses of *Transmission_GPS*, *Detecting_Impact*, and *Send_Message* are to be processed together through *Composite_Action* to locate

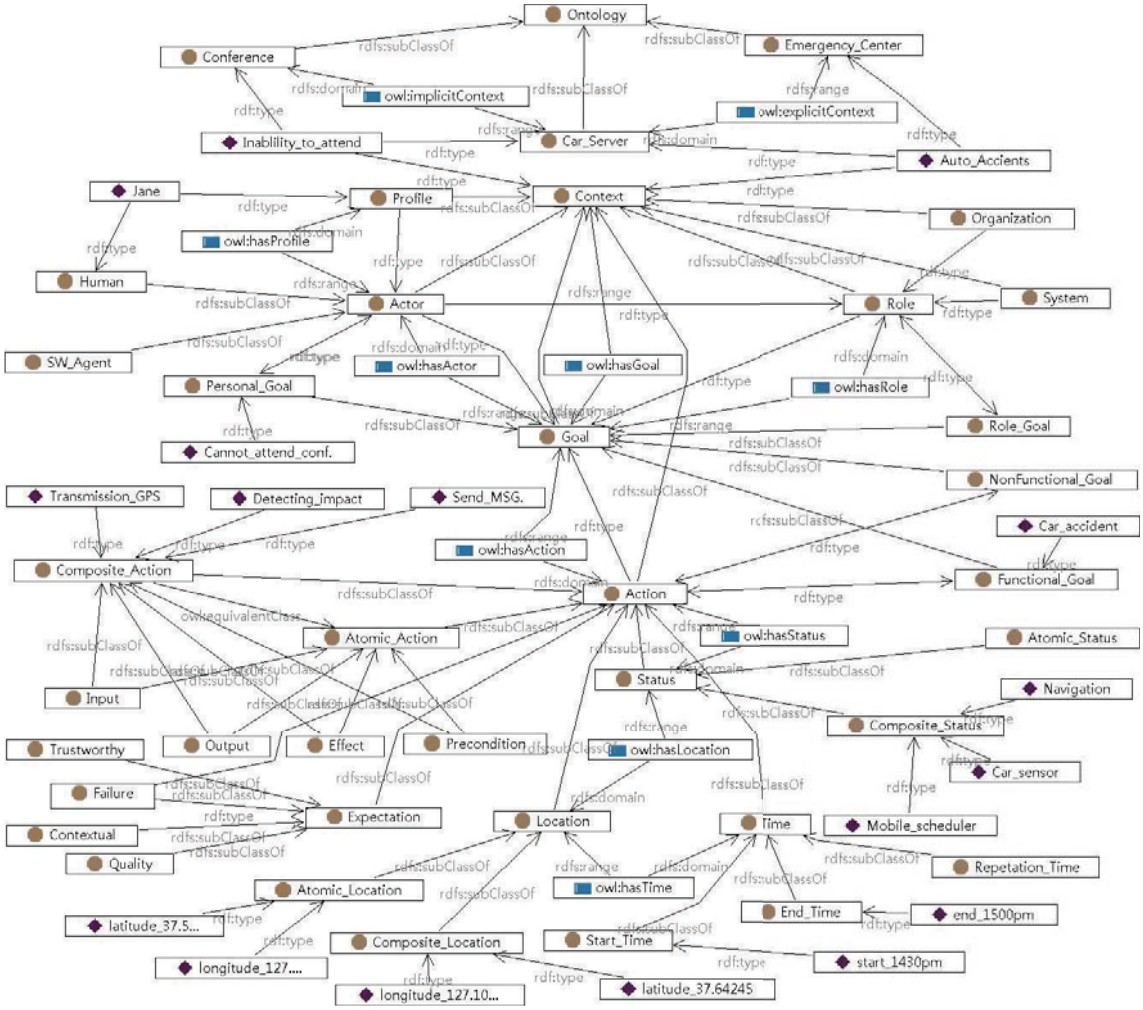


Figure 6: It depicts the classes and properties defined in CA_{5W1H}Onto.

the accident point for processing explicit goal as to the accident, to determine whether or not an accident indeed has occurred, and to determine whether to send an absence message (i.e., implicit goal), respectively. *Composite_Status* has as its instances the navigation to operate `what::Status` and to process explicit goal, the mobile handset (i.e., *mobile_scheduler*) to process *Car_Sensor* with implicit goal. Finally, to process the onset point of the context, the process retains relations with `where::Location` for understanding the location and `when::Time` for understanding the time. `where::Location` has as its subclasses *Atomic_Location* for processing the information on a single location, and *Composite_Location* for processing the information on multiple locations. The location information retains longitudinal and latitudinal coordinates as instances of *Atomic_Location*. `when::Time` processes the information on the time of the accident. It has *Start_Time* and *End_Time* as its subclasses. In the scenario, *2:30 pm* makes the instance of the former, while *3:00 pm* makes that of the latter.

Definition of Scenario-Based Contexts. Model-based CA_{5W1H}Onto

CA_{5W1H}Onto = {**Concepts** → *Car_Server, Emergency_Center, Conference;*

Instances → *Detecting_impact, Transmission_GPS, Cannot_attend_Conf, Car_accident, Jane, Send_message, Car_sensor, Navigation, Mobile_scheduler, Latitude, Longitude, 2:30 PM, 3:00 PM;*

Contexts → *Inability_to_attend, Auto_accident*};

Context Model = {**5W1H**(*What, Where, When, Why, Who, How*)};

Who = {**Role**(*Actor* → *Jane*)};

Why = {**Goal**(*Functional_Goal* → *Car_accident; Personal_Goal* → *Cannot_attend_Conference*)};

How = { **Action** (*Composite_Action* → *Detecting_impact, Transmission_GPS, Send_message*) }

What = {**Status** (*Composite_Activity* → *Car_sensor, Navigation, Mobile_scheduler*)};

Where = {**Location**(*Atomic_Location* → *Latitude, Longitude; Composite_Location* → *Latitude, Longitude*)};

When = { **Time**(*Start_Time* → *2:30 PM; End_Time* → *3:00 PM*) } .

5. Conclusion and Future Work

Under ubiquitous computing circumstances, users and systems should be able to detect and process contextual changes on a real-time basis to provide services appropriate for the changes. Therefore, we employed an ontology-based context-aware modeling technique, based on the maxim of five Ws and one H (i.e., CA_{5W1H}Onto). Then, we proposed a framework enabling efficient specification of contextual information as a method of characterizing and substantializing entities pertaining to the contextual information existing in the ubiquitous computing environment. For that purpose, we also defined and implemented the ontological context-aware schema. The proposed CA_{5W1H}Onto model independently separates ontologies and contexts in the form of modules, prior to defining them. In short, our proposed model is advantageous in terms of adaptability and interoperability with the ontologies already developed in diverse domains. Due to these features, the proposed model shows high levels of expandability and recyclability.

In the near future, we plan to quantitatively evaluate our proposed framework. Furthermore, we intend to study algorithms and rules for context reasoning.

Acknowledgments

This work was supported by the Second Brain Korea 21 Project, research funding from Seoul R&BD Program (WR080951), and the basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education, Science and Technology (2011-0025588). D.-K. Baik is the corresponding author.

References

[1] G. D. Abowd, M. Ebling, G. Hunt, L. Hui, and H. W. Gellersen, "Context-aware computing," *IEEE Pervasive Computing*, vol. 1, no. 3, pp. 22–23, 2002.

[2] B. Schilit, N. Adams, and R. Want, "Context-aware computing applications," in *Proceedings of the IEEE Workshop on Mobile Computing Systems and Applications*, pp. 85–90, Santa Cruz, Calif, USA, December 1994.

[3] A. Schmidt, M. Beigl, and H. W. Gellersen, "There is more to context than location," *Computers and Graphics*, vol. 23, no. 6, pp. 893–901, 1999.

[4] A. Schmidt and K. Van Laerhoven, "How to build smart appliances?" *IEEE Personal Communications*, vol. 8, no. 4, pp. 66–71, 2001.

[5] K. Cheverst, K. Mitchell, and N. Davies, "Design of an object model for a context sensitive tourist GUIDE," *Computers and Graphics*, vol. 23, no. 6, pp. 883–891, 1999.

[6] C. Ghidini and F. Giunchiglia, "Local Models Semantics, or contextual reasoning = locality + compatibility," *Artificial Intelligence*, vol. 127, no. 2, pp. 221–259, 2001.

[7] Web Ontology Language (OWL), http://www.w3.org/2004/OWL/.

[8] Resource description framework (RDF), http://www.w3.org/RDF/.

[9] F. Baader, D. Calvanese, D. McGuinness, D. Nardi, and P. Patel-Schneider, *The Description Logic Handbook*, Cambridge University Press, 2003.

[10] Protégé, RDF/OWL editor, http://protege.stanford.edu/overview/protege-owl.html.

[11] TopBraid Composer, TopQuadrant, http://topquadrant.com/products/TB_Composer.html/.

[12] M. Samulowitz, F. Michahelles, and C. linnhoff-popien, "Capeus: an architecture for context-aware selection and execution of services," in *New Developments in Distributed Applications and Interoperable Systems*, pp. 23–39, Kluwer Academic, 2001.

[13] B. Schilit, N. Adams, and R. Want, "Context-aware computing applications," in *Proceedings of the Workshop on Mobile Computing Systems and Applications*, pp. 85–90, December 1994.

[14] WAPFORUM. User Agent Profile (UAProf), http://www.wapforum.org.

[15] V. Akman and M. Surav, "The use of situation theory in context modeling," *Computational Intelligence*, vol. 13, no. 3, pp. 427–438, 1997.

[16] CoBrA(Context Broker Architecture), http://cobra.umbc.edu/.

[17] A. Salden, R. Poortinga, M. Bouzid et al., "Contextual personalization of a mobile multimodal application," in *Proceedings of the International Conference on Internet Computing (ICOMP '05)*, pp. 294–300, Las Vegas, Nev, USA, June 2005.

[18] T. Gu, H. K. Pung, and D. Q. Zhang, "A middleware for building context-a ware mobile services," in *Proceedings of the 59th IEEE Vehicular Technology Conference (VTC '04)*, pp. 2656–2660, May 2004.

[19] M. Román, C. Hess, R. Cerqueira, A. Ranganathan, R. H. Campbell, and K. Nahrstedt, "A middleware infrastructure for active spaces," *IEEE Pervasive Computing*, vol. 1, no. 4, pp. 74–83, 2002.

[20] D. Salber, A. K. Dey, and G. D. Abowd, "Context toolkit: aiding the development of context-enabled applications," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '99)*, pp. 434–441, Pittsburgh, Pa, USA, May 1999.

[21] X. H. Wang, D. Q. Zhang, T. Gu, and H. K. Pung, "Ontology based context modeling and reasoning using OWL," in *Proceedings of the 2nd IEEE Annual Conference on Pervasive Computing and Communications Workshops (PerCom '04)*, pp. 18–22, March 2004.

[22] DAML+OIL, http://www.daml.org/.

[23] G. D. Abowd, D. Anind, P. J. Brown, N. Davies, M. Smith, and P. Steggles, "Towards a better understanding of context

and context-awareness," in *Proceedings of the 1st International Symposium on Handheld and Ubiquitous Computing (HUC '99)*, pp. 304–307, 1999.

[24] The World Wide Web Consortium (W3C), http://www.w3.org/standards/.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 382316, 7 pages
doi:10.1155/2012/382316

Research Article

A Communication Framework in Multiagent System for Islanded Microgrid

Hak-Man Kim[1] and Yujin Lim[2]

[1] *Department of Electrical Engineering, University of Incheon, 12-1 Songdo-Dong, Yeonsu-Gu, Incheon 406-840, Republic of Korea*
[2] *Department of Information Media, University of Suwon, San 2-2 Wau-Ri, Bongdam-Eup, Hwaseong-Si, Gyeonggi-Do 445-743, Republic of Korea*

Correspondence should be addressed to Yujin Lim, yujin@suwon.ac.kr

Received 6 October 2011; Accepted 19 December 2011

Academic Editor: Tai Hoon Kim

Microgrids are integrated energy systems consisting of interconnected loads and distributed energy sources which as a system can operate in parallel with the grid or in an island mode. By fault occurrence in the connected power system or geographical isolation, microgrids operate in the islanded mode. In the islanded mode, microgrids should be operated to meet a power balance between supply and demand without power trade. In recent years, multiagent systems have been proposed to provide intelligent energy control and management systems in microgrids. In this paper, we design a communication framework to control and operate distributed sources and loads in the islanded microgrids. The framework reliably delivers microgrid control frame between agents by employing wireless mesh network as an advanced topology of the wireless sensor network. From results of experiments, we show that our framework outperforms other conventional one, with respect to the rate of success on the transmission of frames among agents.

1. Introduction

A microgrid is an energy community having clean energy sources such as solar power, wind power, and fuel cells and energy storage devices such as batteries. The energy sources and energy storage devices are distributed in the community, and they are called distributed generation systems (DGs) and distributed energy storage devices (DSs), respectively. Recently, attention on the microgrid has been growing as an eco-friendly power system reducing climate change. Since Professor Lasseter proposed the concept of the microgrid in 2001 [1], many technologies such as power control [2–5], protection schemes [6, 7], simulators, and field tests [8–10] for microgrids have been studied. In addition, multiagent system applications have been studied for efficient and economic control and operation of microgrids [11–20].

The microgrid can be operated by two operation modes: the grid-connected mode and the islanded mode. In the grid-connected mode, a microgrid is connected to a power system, especially a distributed system. On the other hand, the islanded mode means an isolated operation mode from any power system for the case of fault occurrence in the connected power system or geographical isolation such as a small island. In the islanded mode, microgrids should be operated to meet a power balance between supply and demand without power trade. Whenever a power imbalance occurs, the output of DGs is decreased and load shedding is used to solve the power imbalance.

Since DGs and DSs are distributed geographically, microgrids can be operated and controlled using communication links such as the Internet, the power line communication (PLC), and fiber-optic lines [19]. In particular, in the case of geographically islanded microgrids located at an island, the PLC and the wireless sensor network (WSN) can be considered as economical communication links. The PLC is used for the pilot microgrid of Kythnos Island (Greece) [12]. However, the propagation problem and the limited data rates of the PLC are well-known problems. Besides, there

are many ways in which error has been introduced into the communication signals. Interference, cross chatter, some active devices, and some passive devices introduce noise or attenuation into the signal. When error becomes significant, the devices controlled by the unreliable signal may fail, become inoperative, or operate in an undesirable fashion. For this reason, the WSN was considered basically for a communication link in geographically islanded microgrids as explained in our previous work [15].

In this paper, we propose a communication infrastructure based on the WSN for geographically islanded microgrid operated and controlled by a multiagent system. As an advanced topology of the WSN, we employ wireless mesh network (WMN) that needs only a few access points for wireless connections among agents and also reduces the infrastructural costs. To improve the performance in terms of the success on the transmission of frames among agents, we customize the routing protocol for adjusting routes according to the link quality. Also, we verify that our protocol improves the success on the transmission irrespective of dynamics of link quality.

The remainder of this paper is structured as follows. Section 2 describes geographical islanded microgrid operation based on the multiagent system as backgrounds. Section 3 explains our communication infrastructure and discusses the proposed routing protocol. Following this, we verify the designed system by NS-2 simulations in Section 4. Finally, Section 5 summarizes our study results.

2. Islanded Microgrid Operation Based on Multiagent System

2.1. Islanded Microgrid Operation and Control. The microgrid should maintain a constant frequency such as 50 Hz or 60 Hz. In practice, some deviation such as ±0.2 Hz is allowed. The frequency affects a power balance between supply and demand. Figure 1 shows an operation scheme for power balance in the islanded microgrid [17]. The information of power supply, power demand, and status of storages is collected and a condition of power balance is checked. If the power supply is greater than the power demand, DSs are selected for charging. Otherwise, DSs decided to discharge and the load shedding is used.

Figure 2 illustrates the operation procedure of the microgrid simply where an operation plan prepared in the previous interval is implemented in the next interval. The operation is related to planning action for operational intervals and is composed of two steps: planning and implementation [15–19]. In general, the interval period is determined by microgrid operation rules, for example, a few minutes or a few dozen minutes. The planning action as a control reference is established in each interval. And then the control action is followed. Table 1 shows the features of the operation and control.

2.2. Multiagent-Based Islanded Microgrid. An agent is considered as an intelligent agent which senses the changes of environment and acts by its design purpose. A multiagent system is composed of multiple agents. In our previous works [15–19], a multiagent system for microgrid was defined as follows:

$$\mathrm{Ag} = \left\{\mathrm{Ag}_{\mathrm{MGOCC}}, \mathrm{AG}_{\mathrm{DG}}, \mathrm{AG}_{\mathrm{DS}}, \mathrm{AG}_{\mathrm{L}}\right\}, \quad (1)$$

where $\mathrm{Ag}_{\mathrm{MGOCC}}$ is the Microgrid Operation and Control Center (MGOCC) agent, $\mathrm{AG}_{\mathrm{DG}}$ is a set of DG agents ($\mathrm{Ag}_{\mathrm{DG}}$), $\mathrm{AG}_{\mathrm{DG}}$ is a set of storage device agents ($\mathrm{Ag}_{\mathrm{DS}}$), and $\mathrm{AG}_{\mathrm{L}}$is a set of load agents (Ag_{L}). The MOGCC agent manages entire operation and control in the microgrid. Each agent operates and controls its DG, DS, and load. The agents communicate with the agent communication language (ACL) and share their knowledge for cooperation. An example of the message for communication among agents is as follows:

> (<performative>: from <agent name>: to <agent name>: content <OAV type data>),

where OAV (objective attribute values) type data is composed of an object, an attribute of the object, and the value of the attribute. Table 2 is the communication protocols and Tables 3 and 4 are the designed performatives for the protocols [17]. Here, P1 is used for interactions between the MGOCC agents and the DG agents and P2 is used for between the MGOCC agents and the Load/DS agents.

3. WSN-Based Communication Infrastructure

3.1. Design of WMN Structure. As an extension of the WSN, the WMN has been recently developed to provide high-quality services and applications over wireless personal area networks, wireless local area networks, and wireless metropolitan area networks [21]. Its applications and services include wireless home Internet access, community and neighborhood networking, public safety and security surveillance systems, intelligent transportation systems, and emergency and disaster networking. The WMN has a hybrid network infrastructure with a backbone and an access network. It is operated in both ad hoc and infrastructure modes with self-configuration and self-organization capabilities.

The WMN is the ideal solution to provide both indoor and outdoor broadband wireless connectivity in urban, suburban, and rural environments without the need for extremely costly wired network infrastructure [22]. The WMN has been envisioned as the economically viable networking paradigm to build up broadband and large-scale wireless commodity networks. Installing the necessary cabling infrastructure not only slows down implementation but also significantly increases installation cost. Therefore, the wired architecture is costly, unscalable, and slow to deploy. On the other hand, building a mesh wireless backbone enormously reduces the infrastructural cost because the mesh network needs only a few access points for connection. This reduction of network installation cost ensures rapid deployment of a metropolitan broadband network even in rural or scarcely populated urban areas. Thus, we employ the WMN to design communication infrastructure for the multiagent system.

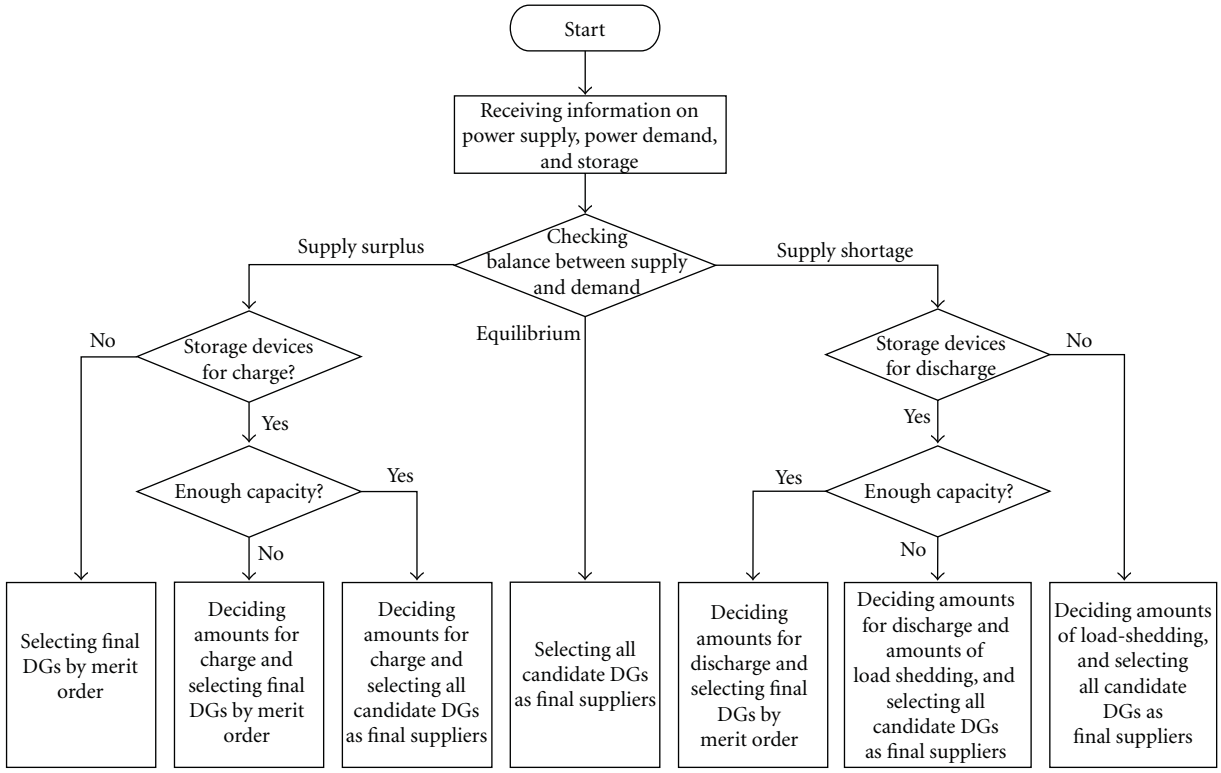


Figure 1: An operation scheme for the islanded operation.

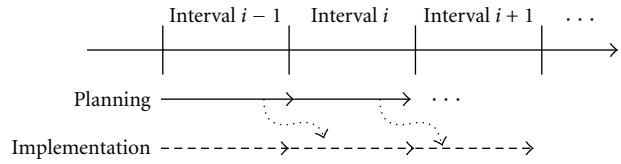


Figure 2: Microgrid operation procedure.

Table 1: The features of operation and control of microgrids.

Action	Interval period	Main function
Operation	(i) A few minutes or a few dozen minutes, for example, 5 min	(i) Establishing operation plans for operation intervals
Control	(ii) A few hundred milliseconds or a few seconds, for example, 1 sec	(ii) Controlling components of a microgrid to follow the operation plan as their control references

Table 2: Agent communication protocols.

Protocols	Executing Agents	Role of Protocol
P1	Between Ag_{MGOCC} and AG_L Between Ag_{MGOCC} and AG_S	Used to exchange data, and to distribute control type and its amount between Ag_{MGOCC} and AG_L, AG_S
P2	Between Ag_{MGOCC} and AG_{DG}	Used to select final suppliers among candidate DGs

The WMN is a group of mesh clients and routers interconnected via wireless links. Mesh clients (MCs) can be various devices with wireless network interface cards such as PCs and laptops. In this paper, a MC represents an agent. Hereafter, MC and agent are exchangeable for convenience. Agents have limited resources and capability in terms of processing ability, radio coverage range, and so on. Mesh routers (MRs) are usually powerful in terms of computation and communication capabilities and have continuous power supply. They normally stay static and act as access points to supply network connections for the agents. Due to limited radio coverage range and dynamic wireless channel capacity, message from an agent usually is transmitted through a multihop path to its destination. Ad hoc mode interconnections of the MRs construct the wireless mesh backbone network. When a new or existing router joins or leaves the backbone, the network self-organizes and self-configures accordingly. In a wireless mesh access network, there are usually one static agent and a number of sensors.

Our WMN structure is illustrated in Figure 3. On each MR, one wireless channels is assigned for access network communication, while the other channel is assigned for the backbone network interconnection. Adjacent access networks should be set to operate on separated channels in order to avoid interference with each other. In the backbone, when directed traffic travels towards the destination, the backbone provides redundant paths between each pair of MRs significantly increasing communications reliability, eliminating single points of failure and potential bottleneck links within

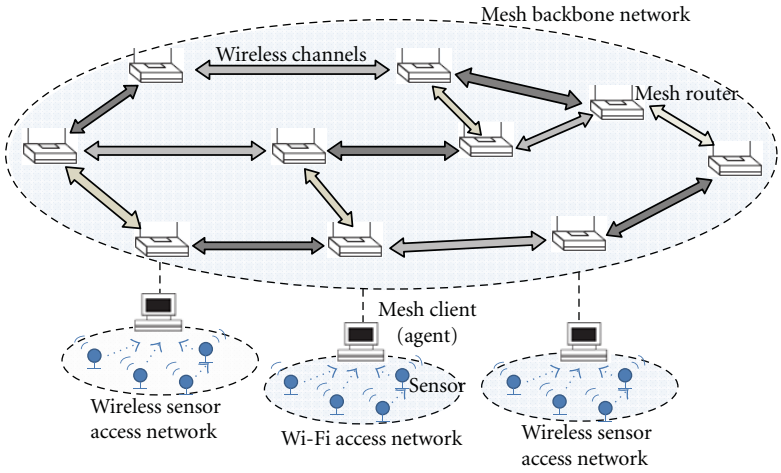


Figure 3: The proposed WMN architecture.

Table 3: Performative for P1.

Performative	Meaning	Remark
Request Information	Request for Information about available storage capacity and charged amount	Between Ag_{MGOCC} and AG_L/AG_S
Receive Informatoin	Receive information	Between Ag_{MGOCC} and AG_L/AG_S
Inform Load	Inform load amount	Between Ag_{MGOCC} and AG_L
Receive Load	Receive load information	Between Ag_{MGOCC} and AG_L
Request Load Shedding	Requests for load shedding	Between Ag_{MGOCC} and AG_L
Inform Storage	Inform available capacity and charged amount	Between Ag_{MGOCC} and AG_S
Receive Storage	Receive storage information	Between Ag_{MGOCC} and AG_S
Request Charge	Request for charge	Between Ag_{MGOCC} and AG_S
Request Discharge	Request for discharge	Between Ag_{MGOCC} and AG_S
Report Load Shedding	Report load shedding	Between Ag_{MGOCC} and AG_L
Report Storage Action	Report action of storage device	Between Ag_{MGOCC} and AG_S

Table 4: Performatives for P2.

Performative	Meaning	Remark
Announce task	Announce to start a new task	Bid price and supply amount
Receive task	Receive a new task	
Bid	Bid for power supply	
Receive Bid	Receive a bid	
Award	Award contracts	
Receive Award	Receive Award	
Report	Report the contract	

the mesh. Network resilience and robustness against potential problems (e.g., node failures, and path failures due to temporary obstacles or external radio interference) are also ensured by the existence of multiple possible routes to the destination.

3.2. Routing Protocol Customized to the Islanded Microgrid. Open standard radio technologies are essential for industry because they bring down the cost of equipment and ensure interoperability. For this reason, several IEEE standard groups are actively working to define specifications for WMN. IEEE 802.11s [23] extends the IEEE 802.11 architecture and protocol for providing the functionality of an extended service set (ESS) mesh. IEEE 802.11s defines a default mandatory routing protocol (Hybrid Wireless Mesh Protocol, or HWMP) [24], yet allows vendors to operate using alternate protocols. HWMP is inspired by a combination of an on-demand AODV [25] and a proactive tree-based routing. The proactive mode requires one MR to be configured as one root MR and we configure the MR connected with MGOCC as the root MR. The root MR constantly propagates routing messages that either establish and maintain routes to all MRs in the mesh or enable MRs to initiate a path to it. In Figure 4, MR K uses the root MR C to establish an initial path (dotted arrow) to MR J. Once

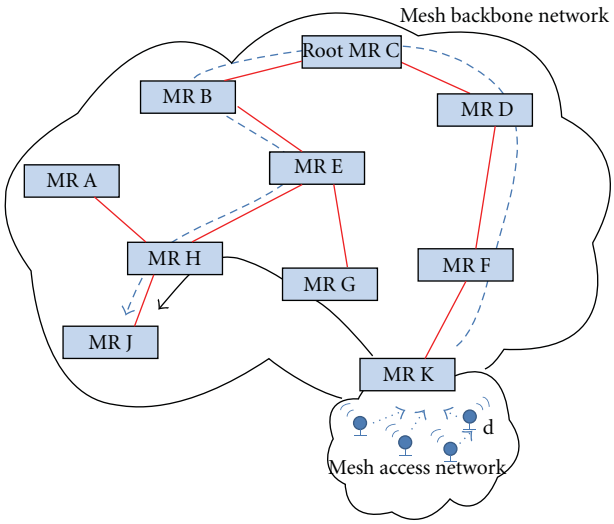


Figure 4: HWMP route selection scheme.

established, MRs use the AODV part of HWMP to avoid the indirection via the root. The K discovers a shorter path (solid arrow) via G and H to forward agent D's frames to the destination MR J.

In on-demand AODV, in order to construct a routing path from source MR to destination MR, the source broadcasts a path request message (PREQ) to the mesh. Upon receiving the PREQ, the destination responds to the source with a path response message (PREP) through unicast transmission. The routing path from the source to the destination is created after the source receives the PREP. Once established, this path is used before it times out. The AODV works on the link layer and adopts a radio-aware metric [26, 27] to calculate the link cost.

The route determined by using the link cost improves the probability of success on the transmission of frames, compared to the route determined by using the hop distance. However, once the route is established, the route is not updated except route failure or route timeout. This route setup procedure decreases the probability of success on the transmission. Since, in our environments, microgrid control messages of agents are exchanged between the MRs, the degradation of the transmission success is fatal. Thus, the dynamic adjusting scheme of routes according to the link quality is needed. However, this scheme suffers from the route flap problem that the MRs constantly change their routes toward the destinations. When the MR detects the degradation of current link quality for the routes, the MR chooses new link having the smallest link cost as its new route. When multiple MRs simultaneously switch to the new link, traffic input rate to the link increases abruptly and congests the link, while the previous link becomes unloaded. Then, the MRs detect the status change and switch back to their previous link. The switch results in unloading the new link, and the process repeats. This route-flap problem causes frequent packet reordering and increases packet loss. We extend HWMP to adjust the route between MRs according to the link quality with no route-flaps. According to the relative link quality, each MR chooses a new link in a statistical manner for delivering the traffic load to its destination.

Table 5: Simulation parameters.

Parameter	Value	Parameter	Value
Slot time	9 μs	Max PREQ retries	3
SIFS	10 μs	Net diameter	35
Signal extension time	6 μs	Diameter traversal time	0.2 s
Preamble duration	20 μs	PREQ min interval	0.1 s
Data rate	54 Mbit/s	PERR min interval	0.1 s
Transmission range	250 m	Path and root timeouts	5 s
Radio-propagation model	Two Ray Ground	Root PREQ interval	1 s

The MR detects the degradation of the link quality of its current link i when the link cost c_i is larger than a certain threshold (c_{th}). Then, the MR selects a new link as follows: the MR calculates $\Gamma(c_i)$ of its links whose c_is are smaller than c_{th}. $\Gamma(c_i)$ is the min-max normalization function of c_i. The normalized link cost represents the relative link quality levels among them. As $\Gamma(c_i)$ of a link is smaller, the link quality of the link is better, compared to the other links having larger values. Given $\Gamma(c_i)$, the MR sorts $\Gamma(c_i)$ in an ascending order and we denote $\{\overrightarrow{\Gamma(c_1)}, \overrightarrow{\Gamma(c_2)}, \ldots, \overrightarrow{\Gamma(c_k)}\}$ as the sorted $\Gamma(c_i)$. Then, with probability p, the MP selects link i as the new link as long as p satisfies the condition as in the following equation:

$$\overrightarrow{\Gamma(c_{i-1})} \leq p < \overrightarrow{\Gamma(c_i)}. \tag{2}$$

4. Performance Evaluations and Discussions

To quantitatively evaluate the performance of the proposed routing protocol, we use NS-2 network simulator [28]. Our WMN delivers the microgrid-related command between agents in a reliable manner. We evaluate the rate of success on the transmission of frames as a performance metric.

In a 1500 m by 1500 m grid, we deploy 20 MCs that is, agents and 36 MRs. The MCs are randomly placed and MRs are equally spaced. The root MR sends a control message every 0.5 seconds. MCs exchange 1024-byte CBR packets with randomly selected other MCs through MRs every second. Table 5 is the list of the parameters as well as the selected values of the physical (PHY), MAC, and HWMP layer applied for the simulation environment. The values for the PHY layer are those for the Extended Rate PHY (IEEE 802.11g). The RTS/CTS mechanism of the MAC layer is disabled by the default settings of the IEEE 802.11 standard. The total simulation time is 3600 seconds.

Figures 5–7 show the probability of success on the frame transmission between MCs. The probability is mostly affected by the frame error rate (i.e., the probability of success on the transmission of frames), data rate, and threshold c_{th} [26, 27]. The numbers in each legend of Figures 5 and 6 present how to set the value of the c_{th}. For example, "Our

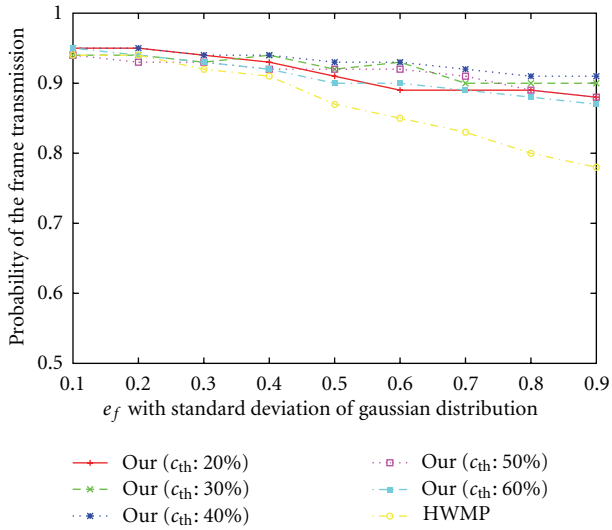


Figure 5: Experimental results obtained by varying frame error rate with standard deviation of Gaussian distribution $N(0, \sigma_1)$.

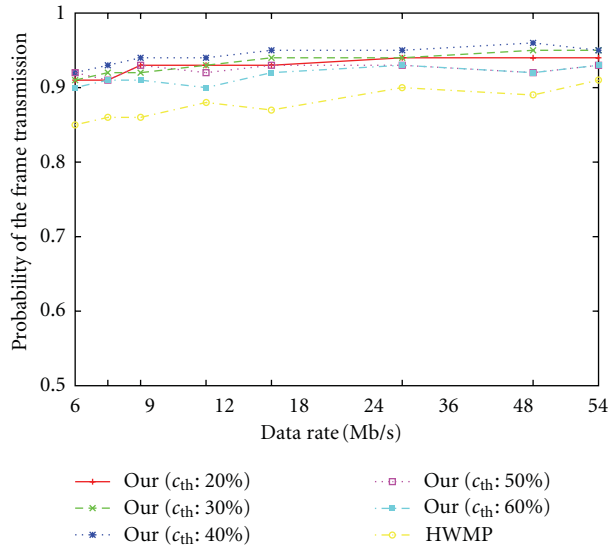


Figure 6: Experimental results obtained by varying data rates.

(c_{th}: 20%)" means our protocol with the c_{th} being 20% of the difference between min and max of the estimated link costs. Since the frame error rate fluctuates in time according to contention and interference, we vary the frame error rate with standard deviation of Gaussian distribution $N(0, \sigma_1)$ in Figure 5. The results indicate that our protocol delivers more frames than HWMP by about 5% on average. When the σ_1 is small, the performances of the both protocols are similar. However, when the σ_1 is large, our protocol delivers more frames up to 11% than HWMP.

Figure 6 shows the performance with varying data rate r when we set σ_1 to 0.3. Since IEEE 802.11 g standard supports data rates of 6, 9, 12, 18, 24, 36, 48, and 54 Mb/s, we vary the data rate according to them. As the data rate increases, the performance usually increases. In our scenario, since

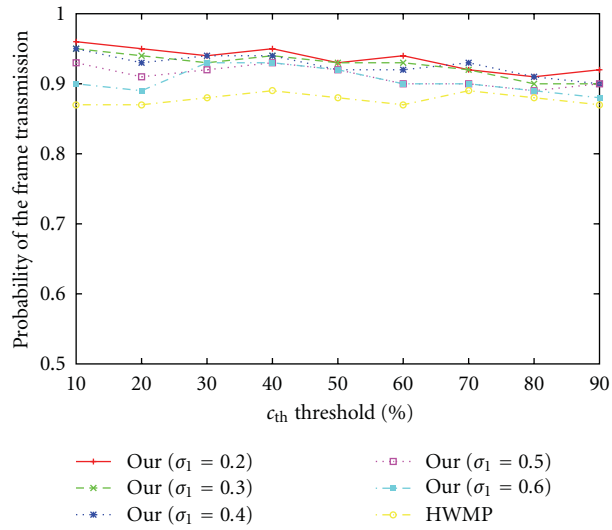


Figure 7: Experimental results obtained by varying c_{th} threshold.

the number of delivered frames between MCs is not large, for example, one frame for each second, the performance improvement is not large. Figure 7 indicates the performance with varying the c_{th}. As seen in the figure, the performance of HWMP is not affected by the c_{th}. In our protocol, as the c_{th} is changed between 10% and 90%, the performance is changed up to 6%. When the c_{th} is small, our protocol is too highly sensitive to the dynamics of the link cost and the overhead of changing the route increases. When the c_{th} is large, the protocol becomes dull and the benefit of adjusting routes according to the link quality decreases. From the results, we conclude that the performance is good when the c_{th} is 30% or 40%.

5. Conclusions

In this paper, we designed a communication infrastructure for multiagent-based islanded microgrid. There are several contributions in our design: we designed the infrastructure to deliver grid control frame between agents by employing WMN. In order to deliver the frame reliably, we extended the traditional HWMP to adjust the route between MRs serving agents according to the link quality. For the development, we selected a new link in a statistical manner to avoid the route-flap problem. In addition, we showed the feasibility of our protocol from the comparison of our protocol with HWMP. From the comparison, we can conclude that the performance of our protocol is better than that of HWMP with respect to the rate of success on the transmission of the frame.

In our protocol, one of the factors affecting the performance is the threshold c_{th}. To optimize the performance, we need any threshold decision scheme in time-varying environments. We will consider developing the scheme as a future research direction. In addition, when the islanded microgrid needs to be connected to the external network, for example, the Internet, our infrastructure uses one or more MRs to connect the external network. As another research direction,

we will develop the load-balancing scheme between the MRs to avoid the traffic concentration.

References

[1] R. H. Lasseter, "Microgrids," in *Proceedings of the IEEE Power Engineering Society Winter Meeting*, pp. 146–149, January 2001.

[2] J. Y. Kim, S. K. Kim, and J. H. Park, "Contribution of an energy storage system for stabilizing a microgrid during islanded operation," *Journal of Electrical Engineering and Technology*, vol. 4, no. 2, pp. 194–200, 2009.

[3] J. Y. Kim, J. H. Jeon, S. K. Kim et al., "Cooperative control strategy of energy storage system and microsources for stabilizing the microgrid during islanded operation," *IEEE Transactions on Power Electronics*, vol. 25, no. 12, Article ID 5580122, pp. 3037–3048, 2010.

[4] C. K. Sao and P. W. Lehn, "Control and power management of converter fed microgrids," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1088–1098, 2008.

[5] R. Majumder, A. Ghosh, G. Ledwich, and F. Zare, "Power management and power flow control with back-to-back converters in a utility connected microgrid," *IEEE Transactions on Power Systems*, vol. 25, no. 2, Article ID 5340642, pp. 821–834, 2010.

[6] H. L. Jou, W. J. Chiang, and J. C. Wu, "A simplified control method for the grid-connected inverter with the function of islanding detection," *IEEE Transactions on Power Electronics*, vol. 23, no. 6, pp. 2775–2783, 2008.

[7] D. M. Vilathgamuwa, P. C. Loh, and Y. Li, "Protection of microgrids during utility voltage sags," *IEEE Transactions on Industrial Electronics*, vol. 53, no. 5, pp. 1427–1436, 2006.

[8] J. H. Jeon, J. Y. Kim, H. M. Kim et al., "Development of hardware in-the-loop simulation system for testing operation and control functions of microgrid," *IEEE Transactions on Power Electronics*, vol. 25, no. 12, Article ID 5594649, pp. 2919–2929, 2010.

[9] Y. Li, D. M. Vilathgamuwa, and P. C. Loh, "Design, analysis, and real-time testing of a controller for multibus microgrid system," *IEEE Transactions on Power Electronics*, vol. 19, no. 5, pp. 1195–1204, 2004.

[10] J.-H. Jeon and S.-K. Kim, "Development of simulator system for microgrids with renewable energy sources," *Journal of Electrical Engineering and Technology*, vol. 1, no. 4, pp. 409–413, 2006.

[11] A. L. Dimeas and N. D. Hatziargyriou, "Operation of a multiagent system for microgrid control," *IEEE Transactions on Power Systems*, vol. 20, no. 3, pp. 1447–1455, 2005.

[12] S. J. Chatzivasiliadis, N. D. Hatziargyriou, and A. L. Dimeas, "Development of an agent based intelligent control system for mierogrids," in *Proceedings of the IEEE Power and Energy Society 2008 General Meeting*, pp. 1–6, July 2008.

[13] A.L. Dimeas and N.D. Hatziargyriou, "Agent based control for microgrids," in *Proceedings of the IEEE Power Energy Society General Meeting*, pp. 1–5, June 2007.

[14] A. L. Dimeas and N. D. Hatziargyriou, "Control agents for real microgrids," in *Proceedings of the 15th International Conference on Intelligent System Applications to Power Systems (ISAP '09)*, pp. 1–5, Curitiba, Brazil, November 2009.

[15] H. M. Kim, T. Kinoshita, Y. Lim, and T. H. Kim, "A bankruptcy problem approach to load-shedding in multiagent-based microgrid operation," *Sensors*, vol. 10, no. 10, pp. 8888–8898, 2010.

[16] H. M. Kim and T. Kinoshita, "A multiagent system for microgrid operation in the grid-interconnected mode," *Journal of Electrical Engineering and Technology*, vol. 5, no. 2, pp. 246–254, 2010.

[17] H. M. Kim, T. Kinoshita, and M. C. Shin, "A multiagent system for autonomous operation of islanded microgrids based on a power market environment," *Energies*, vol. 3, no. 12, pp. 1972–1990, 2010.

[18] H. M. Kim and T. Kinoshita, "A new challenge of microgrid operation," *Communications in Computer and Information Science*, vol. 78, pp. 250–260, 2010.

[19] H. M. Kim and T. Kinoshita, "Multiagent system for Microgrid operation based on power market environment," in *Proceedings of the 31st International Telecommunications Energy Conference (INTELEC '09)*, Incheon, Korea, October 2009.

[20] Z. Xiao, T. Li, M. Huang et al., "Hierarchical MAS based control strategy for microgrid," *Energies*, vol. 3, no. 9, pp. 1622–1638, 2010.

[21] F. Huang, Y. Yang, and L. He, "A flow-based network monitoring framework for wireless mesh networks," *IEEE Wireless Communications*, vol. 14, no. 5, pp. 48–55, 2007.

[22] R. Bruno, M. Conti, and E. Gregori, "Mesh networks: commodity multihop ad hoc networks," *IEEE Communications Magazine*, vol. 43, no. 3, pp. 123–131, 2005.

[23] Draft Amendment to Standard for Information Technology—Telecommunications and Information Exchange Between Systems—LAN/MAN Specific Requirements—part 11: Wireless Medium Access Control (MAC) and physical layer (PHY) specifications: Amendment: ESS Mesh Networking, IEEE P802.11s/D3.0, March 2009.

[24] G. R. Hiertz, D. Denteneer, S. Max et al., "IEEE 802.11s: The WLAN mesh standard," *IEEE Wireless Communications*, vol. 17, no. 1, Article ID 5416357, pp. 104–111, 2010.

[25] C. Perkins, E. Belding-Royer, and S. Das, "Ad hoc on demand distance vector (AODV) routing," IETF RFC 3561, 2003.

[26] Y. D. Lin, S. L. Tsao, S. L. Chang, S. Y. Cheng, and C. Y. Ku, "Design issues and experimental studies of wireless lan mesh," *IEEE Wireless Communications*, vol. 17, no. 2, Article ID 5450658, pp. 32–40, 2010.

[27] R. C. Carrano, L. C. S. Magalhães, D. C. M. Saade, and C. V. N. Albuquerque, "IEEE 802.11s multihop MAC: a tutorial," *IEEE Communications Surveys and Tutorials*, vol. 13, no. 1, Article ID 5473885, pp. 52–67, 2011.

[28] "The network Simulator ns-2," http://www.isi.edu/nsnam/ns/.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 836362, 8 pages
doi:10.1155/2012/836362

Research Article

The Construction of Inference Engine for Meaningful Context and Prediction Based on USN Environment

So-Young Im and Ryum-Duck Oh

Department of Computer Science and Information Engineering, Chung-Ju National University, Chung-Ju, 50 Daehak-ro, Chungbuk 380-702, Republic of Korea

Correspondence should be addressed to Ryum-Duck Oh, rdoh@cjnu.ac.kr

Received 7 September 2011; Accepted 4 December 2011

Academic Editor: Tai Hoon Kim

Currently, with gradually increasing movement to live with nature, artificial wetlands are increasing as well. All these change blows at rivers and streams thereby need for wetland management systems to increase. To measure environmental situations on the wetlands, people should go outside and check with measurement tools regularly. However, with these tools only it is difficult to know the exact situations on that wetland. Thus, we attached various sensors on the wetland and made sensor network environment. We used sensing data from sensor network to assume the situation of the wetland. This paper proposes a design for this through application of context inference of USN (Ubiquitous Sensor Network) and inference production rules for context inference engine of wetland management system by using JESS. In this study, we made rules using actual eutrophication criteria as a standard of water quality. The produced rules in this paper can decide the grade of eutrophication on wetland environment then predict the status of the wetland based on facts collected from sensor networks. Sensors sense data such as DO, BOD, SS, PH. And production rules divided the grades of each fact and then final rules can decide the eutrophication grades which mean water quality grades.

1. Introduction

Wetland is a wet ground surrounded by rivers, ponds, and swamps and also a reserve of natural resources where water always remains under natural environment and lasts long [1, 2]. Thus, ongoing maintenance is required to better manage important wetlands. The wetland needs to be constantly checked in its nature. However, with the use of these measurement tools only, it is difficult to know exactly the situation that wetlands change every time, and the time to prepare may be insufficient.

Accordingly, this study aims to prepare for this through application of context inference techniques of USN (Ubiquitous Sensor Network) and production rules for wetland control system design. The produced rules in this paper can decide the grade of eutrophication on wetland environment then predict the status of the wetland. This paper is organized as follows. Section 2 discusses the basic knowledge on the situation reasoning and rule-based languages needed in designing reasoning rules. And also proposes reasoning engine to predict the state of water quality in the wetland environment. Section 3 applies the performance of the created rules into the wetland environment situation data and then evaluates the performance. Section 4 evaluates the study results and suggests future researches and brings to a conclusion.

2. Related Work

Context inference is an inference of new facts from already-known facts and knowledge. Context inference method is largely divided into two things: rule-based reasoning and case-based reasoning [3]. Rule-based reasoning is a method that operates equipment as long as condition of a sensor is met regardless of place and user. Contrary to this, Case-based reasoning is a method to infer based on the previous cases. Compared to rule-based reasoning, it is easy to learn

```
(name-of-this-production
LHS /*one or more conditions */
–>
RHS /* one or more actions */
)
```

Algorithm 1: A representation method of rules.

and convenient to adapt to exceptional rules. Generally, rule-based inference is applied in an expert system. Rules make it possible to infer another fact based on, the collected data from a sensor (Algorithm 1).

The production memory is a set of productions (rules). A production is specied as a set of conditions, collectively called the left-hand side (LHS), and a set of actions, collectively called the right-hand side (RHS). One condition or more come into LHS, and Action is followed on RHS (reft-hand side), when the conditions are met. Such rules should be set to take a specific action when datum that a sensor collect meets a certain condition or rule that a user has made. Or it can be predicted that a certain situation can happen. This is called a rule-based context inference [4]. For example, it is to perform the specified information when a factor based on water quality data collected from a sensor has a specific data value.

USN means an advanced intelligent community infrastructure which stores, processes, and integrates things and environment information from tags and sensor nodes attached anywhere and can freely use personalized service anytime, anywhere, and anyone through context awareness information and knowledge contents production [5, 6].

Sensor network includes various sensors which can gather information for a specific environment, derive meaningful information based on such collected information, and act like a middleware to lead a role in running appropriate services for users and, finally, obtain information from middleware and interact directly with users (service application). Middleware performs comprehensively: (a) efficiently integrate and control heterogeneous computing resources distributed in Ubiquitous environment; (b) supports integrity; (c) infers the meaning of the various information collected from fundamental roles such as information protection and security functions [7].

The Rete Match Algorithm is a method for comparing a set of patterns to a set of objects in order to determine all the possible matches. It was described in detail in this article because enough evidence has been accumulated to make it clear that it is an efficient algorithm which has many possible applications [8].

JESS is an abbreviation of Java Expert System Shell, and it is a rule-based expert system with features of Java language. JESS is a language which has rule-based system concept similar to LISP but it is easy to define rules and codes and it has a powerful Java API support environment such as networking through integration with Java, graphics, database, and connection [9, 10].

Table 1: Eutrophication criteria.

Eutrophication	Oligotrophic	Mesotrophic	Eutrophic
T-P (mg/m^3)	<15	15~25	>25
Chl-a (mg/m^3)	<3	3~7	>7
SD (m)	>4	4~2.5	<2.5

2.1. Eutrophication. Eutrophication is the addition of artificial or natural substances, such as nitrates and phosphates, through fertilizers or sewage, to an aquatic system [11]. Eutrophication can be human-caused or natural. Untreated sewage effluent and agricultural runoff carrying fertilizers are examples of human-caused eutrophication. However, it also occurs naturally in situations where nutrients accumulate (e.g., depositional environments), or where they flow into systems on an ephemeral basis [12].

Many ecological effects can arise from stimulating primary production, but there are three particularly troubling ecological impacts: decreased biodiversity, changes in species composition and dominance, and toxicity effects [13].

Evaluating the occurrence and progression of eutrophication in the lake is an essential tool for the management of the lake's water quality. The eutrophication evaluation data is the most basic material for the long-term and short-term water management plans. In order to make an accurate assessment of eutrophication in the lake, it is important to collect various measurement data of water quality parameters, but the analysis and interpretation of the collected water quality data is also a very important factor.

The evaluation criteria to determine the eutrophication is divided into (a) evaluation by single-item assessment and (b) evaluation by multiple ones, and so in order to determine the exact status, it is desirable to study with use of synthesized multiple items. With such method, we can estimate the general correlation that exists between water quality items in consideration of physical and chemical characteristics. In addition, if you are using two or more items, you will be able to make more accurate assessment [14].

Table 1 represents the criteria for determining the degree of eutrophication in the lake. The eutrophication is divided into three major steps: we call it an oligotrophic lake when the degree of eutrophication is not severe but, rather, relatively in mild state; a mesotrophic lake when the eutrophication has progressed to some extent; an eutrophic lake when the eutrophication has progressed a lot and needs to be managed.

The items for evaluation are largely divided into three major parts, among these the only item that has the biggest impact on assessment is T-P (total phosphorus). Chl-a is an abbreviation of chlorophyll-a, indicating the concentration of phytoplankton. SD (Secchi depth) is the simplest

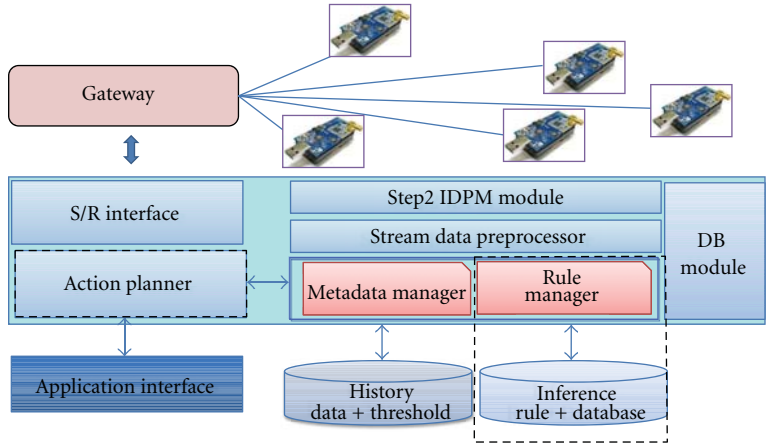


Figure 1: System model of context-aware inference engine.

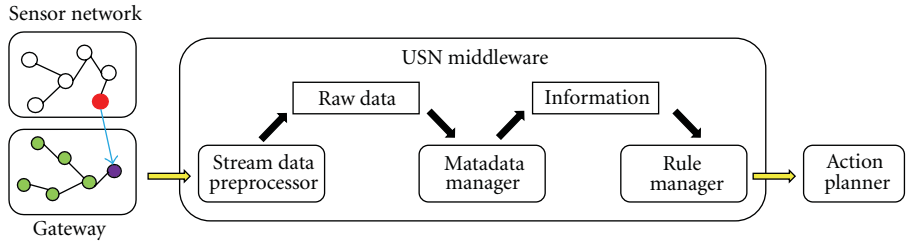


Figure 2: Sensor data flow chart.

indicator to evaluation the eutrophication, being inversely proportional to the concentration of phytoplankton. But in case that suspended matters other than the phytoplankton are present a lot, its disadvantage is that it is less accurate.

3. Configuration of Context Inference System

The main purpose is rule design for context inference engine which allows us to figure out current status of water quality by using JESS language based on rule-based inference method based on collected data from a sensor.

3.1. System Model. Figure 1 represents the System model of context-aware inference engine. System model uses a standard interface for sensor network and consists of incomplete data processing, stream data processing, situation awareness, and reasoning support module.

(1) Interface Manager. As a component which provides a standardized interface for USN sensor networks, each sensor network is a module which complies with common message specifications for sensor data acquisition, monitoring, and control function in the internetworking to sensor network common interface.

(2) Step 2 IDPM (Incomplete Data Processing Mechanism) Module. It is a module which analyzes patterns that appear when an amount of power is insufficient and the patterns of incomplete data resulting from increasing distance of communication distance, determines the relationship between attributes of sensors, and discovers and treats imperfections.

(3) Metadata Manager. It combines packets transmitted from a sensor network with use of metadata and changes it into completed information and reduces frequent access to database with an application of a specific technique, Metadata set.

(4) Rule Manager. It receives environmental information of a sensor network done by Stream Data Processor and matches database to the stored inference rule. Consequently, if necessary, it is reloaded on WM to go through matching tasks, and it comes to perform a particular action depending on reasoning results.

(5) Application Interface. This application interface controls sensor network and provides an appropriate interface with application interface so that it can be monitored.

Figure 2 shows how sensor data flows inside of system. Sensors in sensor network collects an environmental data

Table 2: Kinds of water sensors for eutrophication.

Sensor	Information	Picture
PH	Hydrogen ion exponent	
BOD	Biological oxygen demand	
SS	Suspended solid	
DO	Dissolved oxygen	

and then sends it to the gateway. The stream data goes through the gateway to USN middleware. As figure dedicates, stream data preprocessor receives data packet from gateway then processes raw data out of the stream data. These raw data goes to metadata manager to changes it into completed information for context inference of what situation in such environment. Rule manager take charge of inference tasks to provide to user appropriate service. After matching tasks, the results go to action planner. Action planner does communicate with Application interface to provide inference results to the users.

3.2. Water Sensor Features. Table 2 represents the types and features of sensors necessary to determine the water quality status of the wetlands [15]. PH ranges 6.5~8.5 in any level of water quality in an equal way, and when data outside the range is entered, it is only used to determine that it is abnormal. The other necessary indicators to determine eutrophication are T-N (total nitrogen), T-P (total phosphorus), BOD, SS, DO, and MPN. However, MPN is a measurement of the number of colon bacillus in water and so it is not considered for the design of rules because it cannot be used as a sensor. In addition, as both T-P and T-N are difficult to be measured in a sensor, they are used as a standard to determine the eutrophication using correlations between measurable factors. Here, four facts (BOD, DO, SS, PH) are considered in order to design situation reasoning rules.

3.3. Correlations between Water Quality and Eutrophication. To predict the status of water quality in the wetland using factors that can be measured in a sensor, it is necessary to establish the correlations between such factors and the criteria to determine eutrophication.

To set up such standards, eutrophication assessment criteria and relations commonly used in evaluating the water quality and resources measured from the actual wetland were synthesized and correlation was established. When we make clear the correlation between water quality items and then determine the degree of eutrophication from that, we can estimate the other item out of one item.

Table 3 represents the water quality data is measured on a regular basis in various areas of Lake Chungju. The factors that can be measured in a sensor network are DO, BOD, SS, and PH. As the eutrophication should be determined with the values that can be earned from a sensor, the correlation between these facts was established. Table 4 shows this correlation among the factors.

Table 3: The water quality data of Lake Chungju (2010.08).

No.	DO	BOD	T-N	T-P	SD
1	3.7	1	3.459	0.168	3.5
2	6.7	0.9	2.679	0.007	4.3
3	7.4	7	1	1	2.77
4	8.7	0.9	2.717	0.011	4.7
5	9.6	0.9	2.643	0.008	4.5
6	7.5	0.7	2.632	0.023	3.5
7	11	0.6	2.57	0.006	3.2
8	5.9	0.8	1.836	0.027	2.7
9	8.6	0.9	2.127	0.019	2.6
10	8.9	0.9	2.538	0.018	2.7

Table 4: The correlation.

Correlation
(T-N, T-P) increase → DO decreases → BOD increases
Chl-a decreases → SS decreases
SD decreases → SS increases

DO can be used as one fact of eutrophication criteria, and from which the eutrophication can be determined independently. When you make a reference to the actual data, it shows the signs of increased levels of DO as it transits from being oligotrphic to being eutrophic. BOD is typically used as an indicator to determine water quality, and generally inversely proportional to DO. As the state of the lake transits from being oligotrophic to being eutrophic, the level increases.

SS is a measurement of the amount of suspended solids contained in the lake, being closely related to SD (transparency), which is the criteria to determine eutrophication and inversely proportional. Transparency is closely related to the concentration of N, P, and the concentration of chlorophyll-a changed depending on the growth of algae. But in case of the lake in which there are many other suspended matters other than phytoplankton, it is not a so important indicator for the determination of eutrophication.

PH remains in nearly constant value regardless of water quality. But it is used as criteria to determine whether it is abnormal when levels are too low or too high.

3.4. Context Inference Module and Operating Procedure. Figure 3 shows the operating procedure of inference module. At first a certain event occurs at a network, and data values from a sensor come in (input data). Such raw data values are stored in working memory (WM) set. Users' predefined set of rules are stored in the Data Base. Through comparison process between the sensed values stored in WM and such rules, facts-level value is determined. And then such facts-level values are loaded into WM and are going through matching process with the rules that finally determine water quality.

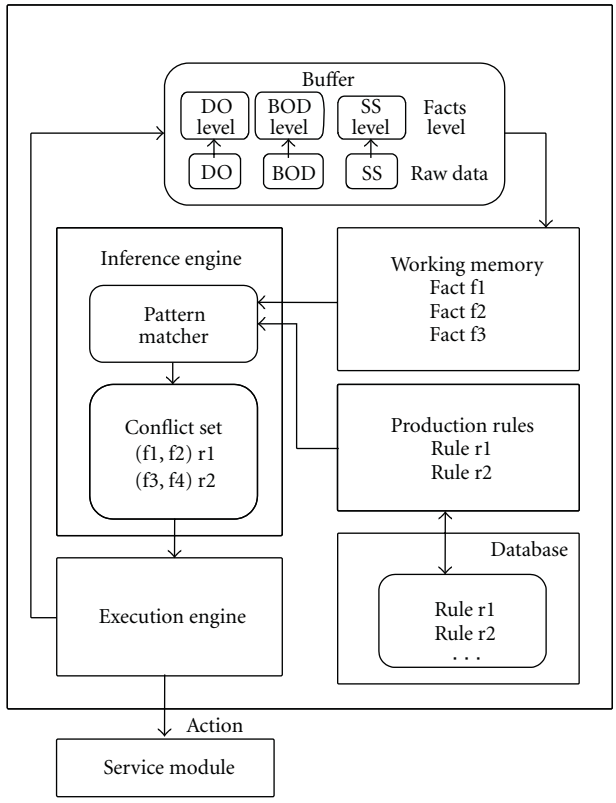


Figure 3: Operating procedure of inference module.

```
(deftemplate IF ‘‘Input-Facts’’;
(slot BOD)
(slot SS)
(slot DO)
(slot PH)
(deftemplate FL ‘‘facts-level’’;
(slot BOD-level)
(slot SS-level)
(slot DO-level)
(slot PH-level))
```

Algorithm 2

```
(defrule DO1
 ?d <- (DO ?x&:(>= ?x 10))
 =>
 (Do-level1)
 (retract ?d)
 (assert (DO-level 1)))
(defrule DO2
 ?d <- (DO ?x&:(and (< ?x 10) (>= ?x 7)))
 =>
 (Do-level2)
 (retract ?d)
 (assert (DO-level 2)))
(defrule DO3
 ?d <- (DO ?x&:(< ?x 7))
 =>
(Do-level3)
 (retract ?d)
 (assert (DO-level 3)))
```

Algorithm 3: The design of rules depending on the correlation with DO. When defining rules at JESS, Defrules are used.

Such process performance is a pattern matcher of inference engine. Pattern Matcher goes through the process of comparison between facts values stored in WM and rules stored in database and stores the matched values into conflict set. The conflict set calculates reliability depending on the order of priority of such matching values and sends matching results that are considered to be the most accurate to execution engine and performs a certain action correspondingly at service module.

3.5. The Design of Production Rule Designs. Rules are designed by using rule-based language, JESS, is a *rule engine* for the Java platform.

JESS can specify logic in the form of rules using the JESS rule language and provide some data for the rules to operate. One of the strengths of JESS is that rules can create new data, or they can do anything that the Java programming language can do. Also JESS that has the advantages is will embed the JESS library in Java code and manipulate it using its own Java API.

Definition of facts. Deftemplate is a function used to define unordered facts in JESS. The facts defined by a function of deftempalte are independent of order if a name is changed by using slot. As the facts that we use here are regardless of order, they are defined as shown in Algorithm 2.

It shows collectible facts from the wetlands environmental sensor networks. Indicators are four kinds of quality management standards. Facts represented as “facts-level” are necessary to design rules to determine the final water quality. Based on water quality management standard, DO (see Algorithm 3), BOD (see Algorithm 4) has 3 standard points and it is divided into 3 stages; SS (see Algorithm 5) into two stages.

Water Quality Standard Rules is based on the relationship with eutrophication criteria to divide DO and BOD into 3 steps and SS into 2 steps.

As DO is a factor that plays a role in determining the eutrophication in the wetland, it is considered with high priority. DO was designed in the way of defining 3 steps of the level. For the rule, it is divided into LHS which belongs to if part on the left and RHS which belongs to then part on the right, on the basis of this symbol of =>.

To take an example of DO1, it is a way of adding fact, “DO-level 1” to WM if the criteria has been met when the value coming from a sensor proves to be 10 (mg/L) or more. Do-level ranges from level 1 to 3, which means that as it is closer to Do-level 1, the level of DO is higher. This also means that it is less likely to be in eutrophication depending on the correlation with the eutrophication criteria. Conversely, if it

```
(defrule BOD1
 ?b <- (BOD ?x&:(>= ?x 1))
 =>
 (Bod-level1)
 (retract ?b)
 (assert (BOD-level 1)))
(defrule BOD2
 ?b <- (BOD ?x&:(and (< ?x 1) (>= ?x 0.7)))
 =>
 (Bod-level2)
 (retract ?b)
 (assert (Bod-level 2)))
(defrule BOD3
 ?b <- (BOD ?x&:(< ?x 0.7))
 =>
(Bod-level3)
 (retract ?b)
 (assert (BOD-level 3)))
```

Algorithm 4: The rules to determine the state of water quality depending on the values of BOD measured.

```
(defrule SS1
  ?s <- (SS ?x&:(>= ?x 1.7))
  =>
  (Ss-level1)
  (retract ?s)
  (assert (SS-level 1)))
(defrule SS2
  ?s <- (SS ?x&:(< ?x 1.7))
  =>
  (Ss-level2)
   (retract ?s)
   (assert (SS-level 2)))
```

Algorithm 5: Standard rules of water quality depending on SS.

```
(defrule PH
  ?ph <- (PH ?x&:(or (> ?x 9) (< ?x 6)))
  =>
  (Ph-level)
  (retract ?ph))
```

Algorithm 6: Detection of Dangerous Situations depending on PH.

is on level 3, it means that it is bad in water quality and that it is likely to be in eutrophication. Therefore, the importance of this was considered on the design of rules and the priority among rules was set higher.

Likewise, DO consists of 3 steps, through which it operates in the way of setting up level depending on the range of BOD level. As we've seen from the correlation above, BOD is inversely proportional to DO. It means that as it is closer to BOD-level 1, BOD value becomes higher. From this we can see that DO level is low and that water quality is not good. Adversely, as it goes closer to level 3, BOD level decreases. And from this, we can presume that water quality is better.

It shows the water quality assessment rules depending on SS values. SS is a factor that is closely related to SD (transparency), one of the eutrophication criteria and regarded as a factor to replace the transparency.

As it is closer to SS-level 2, the amount of suspended solids increases, indicating that its transparency is lower, which in turn means that water quality is not good. But in the case of transparency, if suspended matters other than plankton are plenty in the lake, accuracy of transparency falls down, and its importance is considered as relatively low in the design of rules, and the priority of rules is set in the low.

PH (see Algorithm 6) does not affect to the determination of eutrophication but, if PH is too high or too low, it is possible to determine whether it is in dangerous situation or not, and so warning message is set to be printed when PH is more than 9 or when PH is less than 6.

A part of the final rule which was made by synthesizing BOD, SS, DO level facts. If it's Do-level 2, if BOD-level is less than 2, and if SS-level is 2 or less, it's determined as Grade D (Warning) for warning message to be printed out (see Algorithm 7).

The risk of eutrophication can be divided into six steps: stable (A), to observed (B), cautious (C), warning (D), dangerous (E), and urgent (F). The criteria for determining the risk is set based on the importance of rules. DO is an important factor to determine the status of eutrophication, and the priority is set much higher than the other factors like BOD and SS in setting its rating.

As a rule to determine the ratings of eutrophication, there are 18 rules and from which we can predict the status of wetlands. The level of eutrophication is determined based on

```
 (defglobal ?*grade* = A)
(deffunction gradeD()
  (printout t ‘‘grade D -- warning’’ crlf))
(defrule GRADED1
  ?gradedo <- (DO-level 2)
  ?gradebod <- (BOD-level ?x&:(< ?x 2))
  ?gradess <- (SS-level ?y&:(<= ?y 2))
  =>
  (gradeD)
  (retract ?gradedo)
  (retract ?gradebod)
  (retract ?gradess)
  (bind ?*grade* = D))
```

Algorithm 7

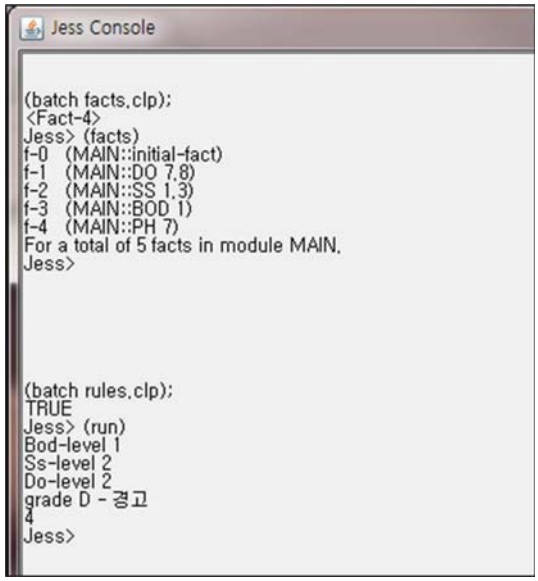


Figure 4: The result obtained after putting in one test data at JESS.

synthesized facts level depending on the coming facts from a sensor network in the wetlands environment.

3.6. Experiment. For design of rules to determine the status of eutrophication in the wetland and tests, data measured from various areas in Lake Chungju is used for experiments. Table 5 indicates 10 randomly selected data among the water quality data in Lake Chungju. An experiment was conducted to figure out if the rules designed by putting in 10 data facts are correctly operated. In order to compare whether the degree of eutrophication is correctly defined, the degree of eutrophication is determined through the eutrophication criteria- T-P, T-N, SD, chl-a- based on test data and then compared. Figure 4 shows the result obtained after putting in one test data at JESS. It is the experimental data having values of DO 7.8, SS 1.3, BOD 1, and PH 7. The results

Table 5: This table indicates 10 randomly selected data among the water quality data in Lake Chungju.

No.	PH	DO	BOD	SS
1	7.6	3.7	1	3
2	8.1	5.8	1.1	2
3	8	6.3	1.1	1
4	7.4	7	1	1
5	8	6.7	0.9	1
6	7.1	8.7	0.9	0.3
7	7.4	9.6	0.9	05
8	7.9	9.9	0.7	0.7
9	7	7.8	1	1.3
10	8.5	11	0.6	2

earned from the rules are Bod-level 1, Ss-level 2, and Do-level 2. From this, we can see that grade D (warning) is set and the warning message is printed out. The degree of eutrophication is determined from the results earned through such reasoning process and the actual data and then compared. Also Table 6 represents the comparative result between the test result and the criteria for the actual degree of eutrophication. We can see that the inference result and the degree of eutrophication in the lake is 80% agreeable. This result shows that it is highly accurate between reasoning result using facts measured from sensor network and the result measured in the actual field.

4. Conclusions

In the past, in order to exactly determine the ratings of water quality in the wetland, for example, streams and rivers, a person has to go out to the field himself or herself on a regular basis in order to measure, compare, and determine.

Accordingly, this paper applies the concept of ubiquitous to the wetland environment, based on facts that are collectible from a sensor network, determines the correlation between eutrophication criteria, that are determined from the actual field, and designs the situation reasoning rules that are available to determine the ratings of eutrophication. From such rules, we can predict the status of water quality in the wetland much faster using reasoning rules without any efforts from a person's hands. This paper uses data collected from the actual wetland environment, applies situation reasoning rules for the designed wetland water quality, and proves its accuracy.

Furthermore, based on such rules, more research is needed to improve accuracy through the designs of various situations that may happen in the wetland and the detailed rules that can be applied to more water quality data. And these rules should be reflected on the design of the situation reasoning engine and the situation reasoning engine should be constantly developed so that a user can obtain more accurate reasoning results.

Table 6: The comparative result.

No.	PH	DO	BOD	SS	Test Result	Eutrophication	Match
1	7.6	3.7	1	3	F	E, F	O
2	8.1	5.8	1.1	2	F	E, F	O
3	8	6.3	1.1	1	E	E, F	O
4	7.4	7	1	1	C	C, D	O
5	8	6.7	0.9	1	E	A, B	X
6	7.1	8.7	0.9	0.3	C	C, D	O
7	7.4	9.6	0.9	0.5	C	A, B	X
8	7.9	9.9	0.7	0.7	B	A, B	O
9	7	7.8	1	1.3	D	C, D	O
10	8.5	11	0.6	2	A	A, B	O

Acknowledgments

This research was financially supported by the Ministry of Education, Science Technology (MEST) and Korea Institute for Advancement of Technology (KIAT) through the Human Resource Training Project for Regional Innovation and the research was supported by a grant from the Academic Research Program of chungju National University in 2011.

References

[1] Water Environmental Kit, http://www.kwater.or.kr/.
[2] Natlonal Institute of Environmental Research Environ mental Kit, 1990, http://library.nier.go.kr/.
[3] L. Hyun-soo and L. Ju-hyun, "A Study on the conceptual model of the ubiquitous residential environment based on the context-aware inference," September 2009.
[4] R. B. Doorenbos, "Production matching for large learning systems," Computer Science Department, Carnegie Mellon University, 1995.
[5] USN Concept, structure, http://ain.knu.ac.kr/professor/APNOMS2005_Tutorial_jtpark_v3.pdf.
[6] U-Sensor Network, ETRI, 10.2004, http://www.etri.re.kr/eng/.
[7] H. Yong Youn and W. D. Cho, "U-computing middleware technology," June 2010.
[8] C. L. Forgy, "Rete: a fast algorithm for the many pattern/many object pattern match problem," *Artificial Intelligence*, vol. 19, no. 1, pp. 17–37, 1982.
[9] "Jess, the rule engine for the Java platform," http://www.jess-rules.com/jess/docs/index.shtml.
[10] E. J. Friedman-Hill, "Jess, The Java Expert System Shell," December 1998.
[11] *Over Fertilization of the World's Freshwaters and Estuaries* , University of Alberta Press.
[12] J. Bartram, W. W. Carmichael, I. Chorus, G. Jones, and O. M. Skulberg, "Chapter 1. Introduction," in *Toxic Cyanobacteria in Water: A Guide to Their Public Health Consequences, Monitoring and Management*, World Health Organization, 1999.
[13] L. Horrigan, R. S. Lawrence, and P. Walker, "How sustainable agriculture can address the environmental and human health harms of industrial agriculture," *Environmental Health Perspectives*, vol. 110, no. 5, pp. 445–456, 2002.
[14] Sun-Hwan Park, Report for water quality of Artificial lake, 2003.
[15] DreamTest, "Kinds of sensors," http://www.dreamtest.co.kr/.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 819497, 7 pages
doi:10.1155/2012/819497

Research Article

Grid-Based Predictive Geographical Routing for Inter-Vehicle Communication in Urban Areas

Si-Ho Cha,[1] Keun-Wang Lee,[1] and Hyun-Seob Cho[2]

[1] *Department of Multimedia Science, Chungwoon University, Chungnam 350-701, Republic of Korea*
[2] *Department of Electronic Engineering, Chungwoon University, Chungnam 350-701, Republic of Korea*

Correspondence should be addressed to Keun-Wang Lee, kwlee@chungwoon.ac.kr

Received 22 October 2011; Accepted 2 January 2012

Academic Editor: Tai Hoon Kim

Vehicular ad-hoc networks (VANETs) are highly mobile wireless ad hoc networks for vehicular safety and other commercial applications, whereby vehicles move non-randomly along roads while exchanging information with other vehicles and roadside infrastructures. Inter-vehicle communication (IVC) is achieved wirelessly using multihop communication, without access to fixed infrastructure. Rapid movement and frequent topology changes cause repeated link breakages, increasing the packet loss rate. Geographical routing protocols are suitable for VANETs. However, they select the node nearest to the destination node as a relay node within the transmission range, increasing the possibility of a local maximum and link loss because of high mobility and urban road characteristics. We propose a grid-based predictive geographical routing (GPGR) protocol, which overcomes these problems. GPGR uses map data to generate a road grid and to predict the moving position during the relay node selection process. GPGR divides roads into two-dimensional road grids and considers every possible node movement. By restricting the position prediction in the road grid sequence, GPGR can predict the next position of nodes and select the optimal relay node. Simulation results using ns-2 demonstrated performance improvements in terms of local maximum probability, packet delivery rate, and link breakage rate.

1. Introduction

VANET is a research field that is attracting growing attention. VANET provides both vehicle-to-infrastructure (V2I) communication and vehicle-to-vehicle (V2V) communication. V2I can provide real-time information on the road traffic conditions, weather, and a basic Internet service via communication with backbone networks. V2V can be used for to providing information about traffic conditions and/or vehicle accidents based on wireless inter-vehicle communication (IVC). In V2V communication environments, vehicles are wirelessly connected by multihop communication without access to some any fixed infrastructure [1]. Already, automobile manufacturers and research centers are investigating the development of IVC protocols for the establishment of VANETs, which are expected to be useful for road safety and many commercial applications [2]. VANET has unique characteristics compared with MANET such as high node mobility and a rapidly changing network topology compared to mobile ad hoc network (MANET). Current VANET routing protocols typically use relay nodes to forward data packets to their destination. However, the rapid movement of vehicles and the frequent topology change of vehicles mean that link breakages occur repeatedly, as shown in Figure 1. Frequent link disconnection is also caused due to other characteristics of VANET such as vehicle movements that are constrained by roads and traffic lights that have greatly affected vehicle movement [3]. The frequent link disconnections may increase the possibility of a local maximum. Because of these problems, geographical routing protocols such as GPSR [4] are known to be more suitable and useful for VANET than existing routing protocols designed for MANETs. For instance, Figure 1 shows that geographical forwarding could use node N2 instead of N1 to forward data to D.

Geographical forwarding is one of the best solutions for VANET routing because it maintains only local information of neighbors rather than per-destination routing entries.

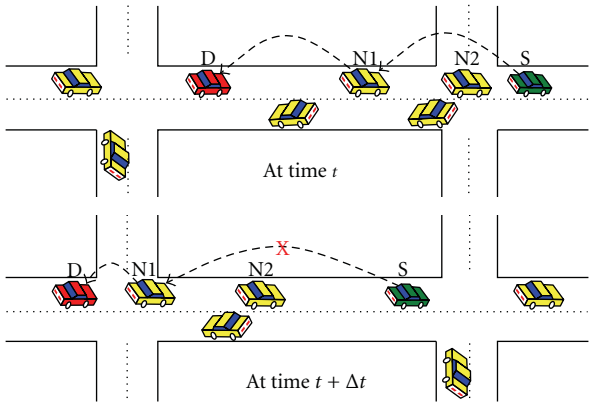


Figure 1: Route (S, N1, D) that was established at time t breaks at time $t + \Delta t$ when N1 moves out of the transmission range of S.

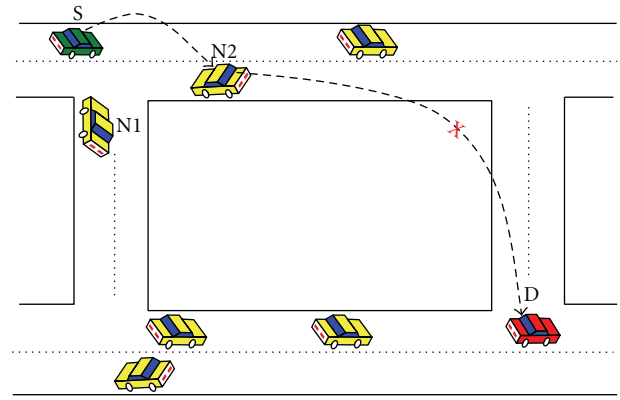


Figure 2: GPSR selects N2 rather than N1 for forwarding the packet, because N2 is closer to D and within the transmission range of S.

GPSR selects the node closest to the destination as the relay node from neighbor nodes. However, GPSR may lead to a link loss problem in urban environments. GPSR does not take into account road structure or the speed and movement direction of vehicles, so it may select stale nodes as relay nodes. Vehicle movements are constrained by roads, so GPSR that fails to consider urban environment characteristics is not suitable for VANET [3]. To solve this problem, greedy perimeter coordinator routing (GPCR) [5] and greedy perimeter urban routing (GPUR) [6] have been proposed as possible solutions. However, GPCR may cause transmission delays and path selection errors because it identifies nodes on a junction by detecting coordinator nodes when selecting relay nodes. GPUR selects nodes with 2-hop neighbors as relay nodes. This causes serious transmission delays. GPUR cannot resolve the local maximum problem because it does not consider road specifications such as dead end streets.

This paper proposes GPGR, which is a grid-based predictive geographical routing protocol for IVC. The protocol uses map data to generate road grids for the path of moving vehicles, and it predicts the exact movement position along the road grids. To do this, we assume that each vehicle knows its location by GPS, as with most related geographic routing protocols, and it uses a grid-based street map for road information. A grid sequence on a road path is a route where vehicles can move. Vehicles can be located on one space of the grid sequence at a specific time. Our target was to improve the routing protocol for IVC, based on vehicle movement information including the position, direction, and velocity on the grid sequence. Position prediction based on a road grid is more realistic rather than blindly predicting whether roads in a segment of roads contain many curves.

The rest of the paper is organized as follows. Section 2 discusses related work while Section 3 introduces the proposed IVC routing protocol. The performance evaluation is presented in Section 4. Finally, our conclusions and future research are outlined in Section 5.

2. Related Work

Traditional MANET routing protocols, such as AODV [7] and DSR [8], are not suitable for VANET, because VANET has unique characteristics such as high node mobility and a rapidly changing network topology compared with MANET. To deal with the rapidly changing network topology of VANET, greedy forwarding protocols have been proposed that are based on geographic information.

GPSR [3] is a well-known greedy forwarding protocol. GPSR makes greedy forwarding decisions using only information about immediate neighbors in the network topology. GPSR may increase the possibility of a local maximum and link breakage, because of the high mobility of vehicles and the specific characteristics of roads in urban areas. This is because it simply selects the nearest node within the transmission range of the destination as a relay node when making packet forwarding decisions. GPSR may also lead to a link loss problem because it maintains stale nodes as neighbor nodes when selecting a relay node in the greedy mode. These local maximum and link breakage problems can be recovered by perimeter mode forwarding, but packet loss and delay time may result because the number of hops is increased by perimeter mode forwarding. This decreases the reliability of VANET. Figure 2 shows an example. Assume that vehicle S wants to send a packet to D, and S has two neighbors: N1 and N2. GPSR will select N2 to forward the packet because N2 is closer to D. However, common sense dictates that we should choose N1 because vehicle movements are constrained by roads.

GPCR [9] was proposed to improve the reliability of GPSR with VANET. The basic behavior of GPCR is similar to GPSR, but it selects a relay node by considering information related to the road structure. GPCR makes routing decisions on the basis of streets and junctions rather than individual nodes and their connectivity. However, GPCR forwards data packets based on the node density of adjacent roads and the connectivity to the destination. Thus, if the density of nodes is low or if there is no connectivity to the destination, the delay time increases and the local maximum problem

is still unresolved. GPUR [5] selects a relay node based on information about the road characteristics, which is similar to GPCR. However, unlike GPCR, GPUR selects a relay node from nodes with 2-hop neighbors. It transmits periodic beacon messages to estimate the presence of 2-hop neighbors among all the relay candidates. The periodic beacon messages that are used to evaluate the presence of 2-hop neighbors lead to serious transmission delays. GPUR also fails to resolve the local maximum problem because it does not consider road specifications such as dead ends.

GSR [10] uses a map and a position-based addressing scheme when sending packets to destinations. The source node evaluates the shortest path between itself and the destination. GVGrid [11] is a source-routing protocol that is similar to GSR. GVGrid finds a network route by route discovery that is expected to provide the best stability, based on a digital map and the positional information of each vehicle. HarpiaGrid [2] is a geography-aware grid-based routing protocol, which uses map data to generate the shortest transmission grid route. This method effectively trades off route discovery communication overheads with insignificant computation time. By restricting the packets in grid sequences rather than a blind greedy search and by making use of a route cache approach, HarpiaGrid reduces many unnecessary transmissions. However, GSR, GVGrid, and HarpiaGrid are all proactive routing protocols. In proactive routing protocols, all vehicles need to maintain a consistent view of the network topology. When a network topology change occurs, the respective updates must be propagated throughout the network to notify the change. Using proactive routing algorithms, vehicles proactively update the network state and maintain a route, regardless of whether data traffic exists, and the overheads of maintaining up-to-date network topology information is high. Thus, they are not suitable for VANETs.

3. Grid-Based Predictive Geographical Routing (GPGR)

GPGR employs road segments based on a routing approach with street awareness, and it uses knowledge of the road topology provided by a static street map. Therefore, data packets will be routed between vehicles, following the road topology and the road segments in the real area. This method aims to improve the routing protocol for IVC based on vehicle movement information such as position, direction, and velocity, as well as the road topology. To do this, we assume that each vehicle knows its location by GPS, was with most related geographic routing protocols, and has a digital street map for road information. Table 1 lists the symbols used in the proposed GPGR algorithm.

The geographic area of VANET is partitioned into a two-dimensional logical grid. Grids are numbered (x, y) following conventional x, y coordinates. Each grid is a sequence area of size $d \times d$ as shown in Figure 3.

Given any physical location, there should be a predefined mapping from the location to its grid coordinates. Where each vehicle has a radio range of r, each grid size d is determined by $d = r/2\sqrt{2}$ to represent the maximum value of d such that a vehicle located at a position in a grid is capable of transmitting data to any vehicle in its eight neighboring grids, as shown in Figure 4.

Table 1: Symbols used for defining GPGR.

Symbols	Definitions
r	Radio range of each vehicle
d	Each grid size
$G(x_i, y_i)$	A grid coordinate for position (x_i, y_i)
V_S	A source vehicle
V_D	A destination vehicle
V_R	A relay vehicle
V_N	A set of neighbor vehicles within a sender's r
$D(V_S, V_N)$	Distance between V_S and V_D
G	A grid sequence of a vehicle
V	Velocity of a vehicle
θ	Moving direction of a vehicle
t	Current time
Δt	Difference in time measurement

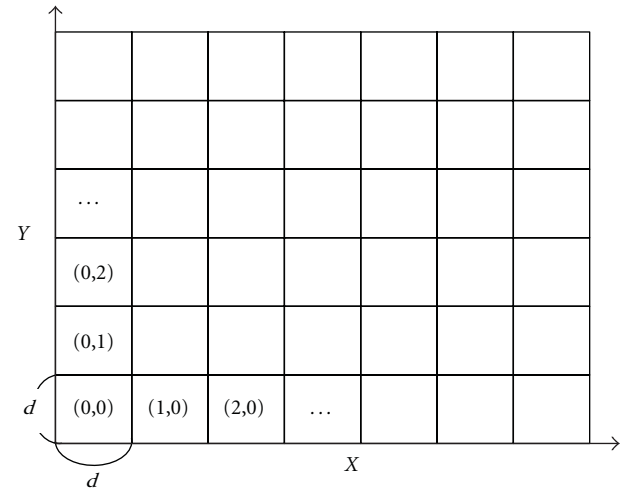


Figure 3: GPGR uses logical grids to partition a physical space.

Our beacon messages contain grid coordinates rather than the positions of vehicles. Vehicles know their own position (x_i, y_i), so each vehicle can calculate its current grid coordinates $G(x_i, y_i)$ based on the floor function given in (1)

$$G(x_i, y_i) = \left[\left\lfloor \frac{x_i}{d} \right\rfloor, \left\lfloor \frac{y_i}{d} \right\rfloor\right]. \quad (1)$$

The procedure for selecting a relay vehicle among all relay candidates is shown in Algorithm 1. It is assumed that sender V_S is sending message MSG to the destination vehicle V_D.

In GPGR, when a source vehicle, V_S, wants to send a message to a destination vehicle, V_D, V_S should first inspect all relay candidates that are similar to GPUR. From these relay candidates, it then selects the node nearest to the destination as a relay node within its transmission range r

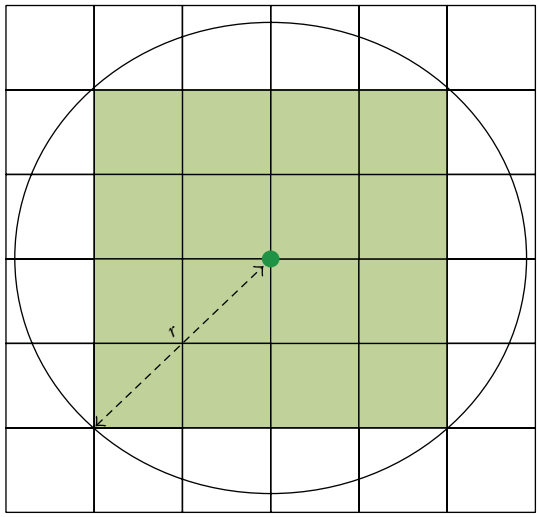


Figure 4: The side length of grids d is determined by $d = r/2\sqrt{2}$ where each vehicle has a radio range of r.

based on the future vehicle position using the road grid. V_S finds all neighbor vehicles $V_N = \{V_j, V_k, \ldots, V_l\}$ in its r and calculates the distance $D(V_S, V_N)$ to find a relay vehicle V_R whose distance is maximal based on the road grid.

For the distance between two grids, GPGR calculates the Euclidean distance between the two centers of the grid cells according to (2)

$$\overline{G(x_i, y_i)G(x_j, y_j)} = \sqrt{(x_i - y_i)^2 + \left(x_j - y_j\right)^2} \times d. \quad (2)$$

An advantage of RPGR is that the predicted positions based on the road grid are more realistic. If a vehicle is on a curved road with no road grid, its position prediction will be incorrect.

Let there be N vehicles present in a road segment space. The location of each vehicle at time t is given by $G(x_i(t), y_i(t))$, where $i = 1$ to N. The next location of each object at time $(t + \Delta t)$ is predicted to be $G(x_i(t + \Delta t), y_i(t + \Delta t))$. Therefore, based on the previous position $G(x_i(t - \Delta t), y_i(t - \Delta t))$ and the current position $G(x_i(t), y_i(t))$, GPGR can predict the next position of the relay candidate at $t + \Delta t$ as $G(x_i(t + \Delta t), y_i(t + \Delta t))$ by using the velocity and direction of the relay candidate.

The velocity and the direction of the relay candidate are given by (3) and (4), respectively

$$V = \frac{\sqrt{(x_i(t) - x_i(t - \Delta t))^2 + (y_i(t) - y_i(t - \Delta t))^2}}{t - (t - \Delta t)}, \quad (3)$$

$$\theta = \tan^{-1}\left(\frac{y_i(t) - y_i(t - \Delta t)}{x_i(t) - x_i(t - \Delta t)}\right). \quad (4)$$

Table 2: Simulation parameters.

Parameter	Value
Topology size	700 × 1000 m
Transmission range	125 m
MAC protocol	IEEE 802.11
Node number	100 to 200
Node velocity	0 km/h to 80 km/h
Beacon time	1 s
Bandwidth	2 Mbps
Packet size	1000 bytes

Therefore, GPGR can calculate the next position of the relay candidate using (5).

$$G(x_i(t + \Delta t), y_i(t + \Delta t)) = \left[\left\lfloor \frac{(x_i(t) + V \times \cos\theta \times \Delta t)}{d} \right\rfloor, \left\lfloor \frac{(y_i(t) + V \times \sin\theta \times \Delta t)}{d} \right\rfloor\right]. \quad (5)$$

If $G(x_i(t + \Delta t), y_i(t + \Delta t))$, the predicted next position of the relay candidate is not on the grid sequence of the road grid, so GPGR should select the closest grid to $G(x_i(t + \Delta t), y_i(t + \Delta t))$ instead of the predicted grid. The grid sequence of the road grid $G = \{G(x_1, y_1), G(x_2, y_2), \ldots, G(x_k, y_k\}$ is generated based on the available map information, where $G(x_2, y_2), \ldots, G(x_{k-1}, y_{k-1})$ are grids in the road path from $G(x_1, y_1)$ to $G(x_k, y_k)$. For example, the grid sequence of the road for moving the vehicles on the road is $\{G(4, 3), G(4, 4), G(4, 5), \{G(3, 5) \mid G(5, 5)\}\}$, as shown in Figure 5. Therefore, all nodes can move along only one of the following grid sequences: $\{G(4, 3), G(4, 4), G(4, 5), G(3, 5)\}$ or $\{G(4, 3), G(4, 4), G(4, 5), G(5, 5)\}$. As shown in Figure 5, the predicted position of the node will be $G(3, 5)$ instead of $G(4, 5)$ or $G(5, 5)$ at $t + \Delta t$ based on the movement direction of the vehicle given by (5). Therefore, the position of the relay candidate can be predicted. The position prediction based on road girds is more realistic if two roads are superposed with or without different running directions.

4. Performance Evaluation

We analyzed and compared the performance of the proposed GPGR and the existing GPSR, GPCR, and GPUR using the ns-2 simulator. In this performance evaluation, we simply considered the local maximum probability, packet delivery rate, and the link breakage rate at this point. The simulated area was based on a real map of Seoul with a 700 × 1000 m size. Table 2 summarizes our simulation parameters. The simulations were performed for 180 s, and the number of nodes was increased from 100 to 200. The movement velocity of the nodes was increased from 0 km/h to 80 km/h. The experiments were performed three times, and the average values were used. Maximum and minimum values were excluded.

/* When V_S sends MSG to V_D */
(1) All V_S get their G to V_D,
$G = \{G(x_1, y_1), G(x_2, y_2), \ldots, G(x_k, y_k)\}$
(2) **For** each V_S **do**
(3) V_S finds all V_N,
$V_N = \{V_j, V_k, \ldots, V_l\}$ in its r.
(4) V_S predicts $G(x(t + \Delta t), y(t + \Delta t))$ of V_N at $t + \Delta t$.
(5) **If** $(G(x(t + \Delta t), y(t + \Delta t)) \notin G)$
(6) V_S finds the nearest alternative grid, $G(x_i, y_i) \in G$ instead of $G(x(t + \Delta t), y(t + \Delta t))$.
(7) V_S replaces $G(x(t + \Delta t), y(t + \Delta t))$ with $G(x_i, y_i)$ such that V_N is on the road.
(8) **End If**
(9) V_S chooses V_R in V_N such that $D(V_R, V_D)$ is minimal.
(10) V_S forwards packets to V_R.
(11) $V_S \leftarrow V_R$.
(12) **End For**

Algorithm 1: Relay vehicle selection.

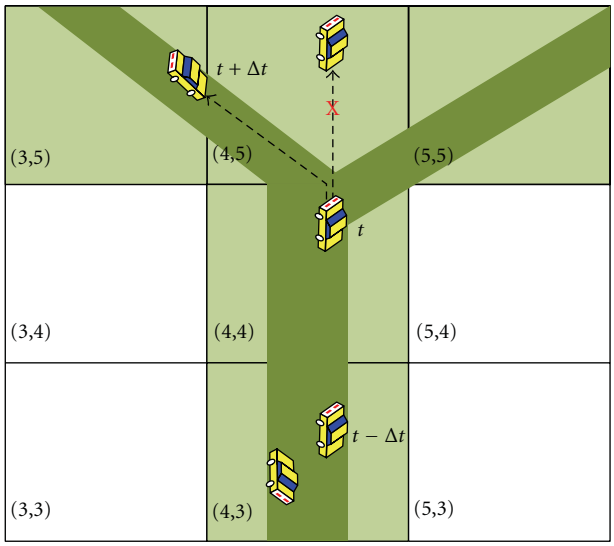


Figure 5: GPGR can predict the next position of nodes and select the optimal relay node by restricting position prediction to the sequence of the road grid.

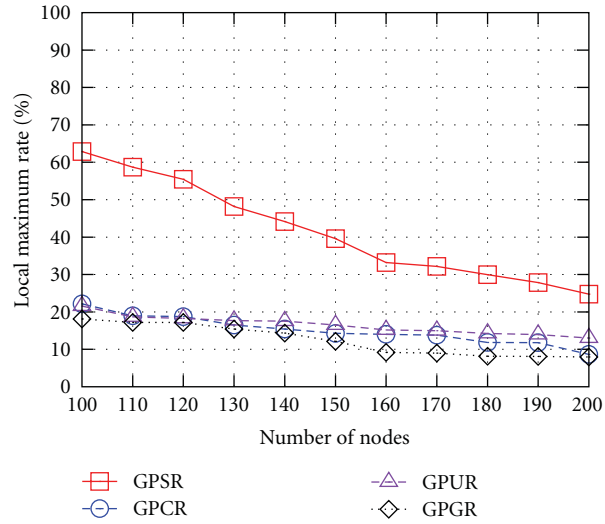


Figure 6: Local maximum rate according to number of nodes.

Figures 6 and 7 show the probability of the local maximum based on the number of nodes and the probability of the local maximum relative to the velocity variation of the nodes, respectively. We observed that a larger number of nodes reduced the probability of the local maximum with all the routing algorithms, as shown in Figure 6. GPGR reduced the probability of the local maximum compared with GPSR, GPCR, and GPUR. This was because GPGR predicted the position of nodes and selected the relay node using a grid map based on the road topology. Therefore, GPGR significantly reduced the probability of a local maximum even when the road topology was changed. GPSR had the highest probability of a local maximum because it selected the nearest node to the destination node as the relay node. GPCR and GPUR had a lower probability of a local maximum compared with GPSR. However, because they selected the relay node based only on the intersection and whether the next relay candidate's neighbor nodes were present or not, they had a higher probability of a local maximum than GPGR. Figure 7 shows that a higher velocity for the nodes led to a higher probability of a local maximum with GPSR, GPCR, and GPUR. However, the probability of a local maximum with GPGR was almost constant, regardless of the velocity of nodes. GPGR reduced the local maximum rate compared with GPSR, GPCR, and GPUR because GPGR can reduce errors by selecting a stale node as a relay node even if the velocity of nodes is increased. GPGR predicts the

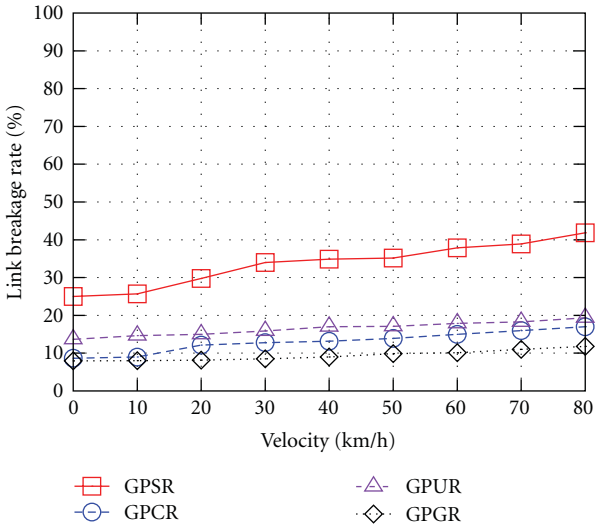


Figure 7: Local maximum rate according to velocity of nodes.

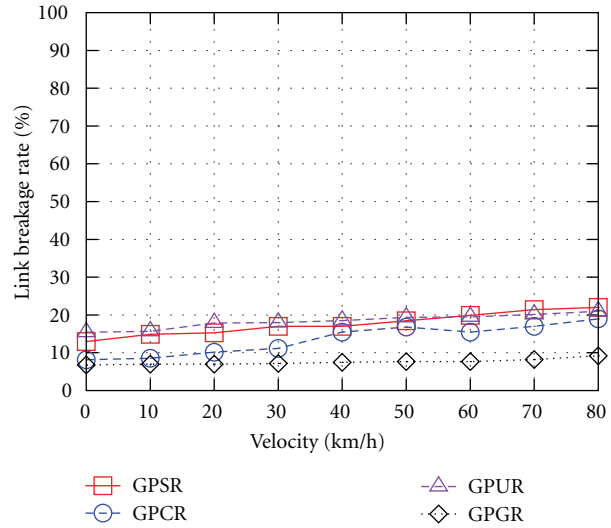


Figure 9: Link breakage rate according to the velocity of nodes.

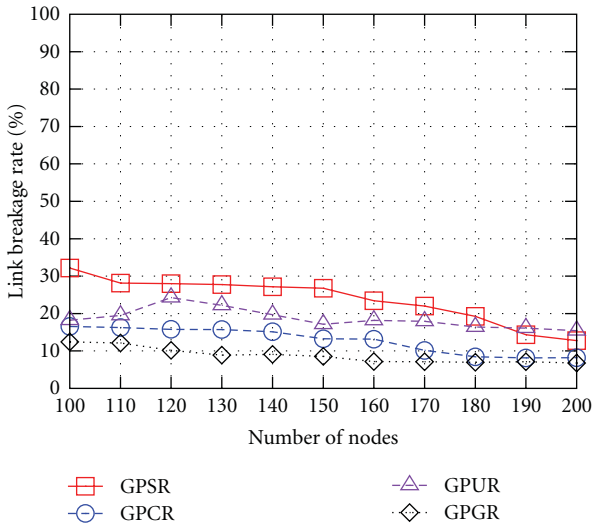


Figure 8: Link breakage rate according to the number of nodes.

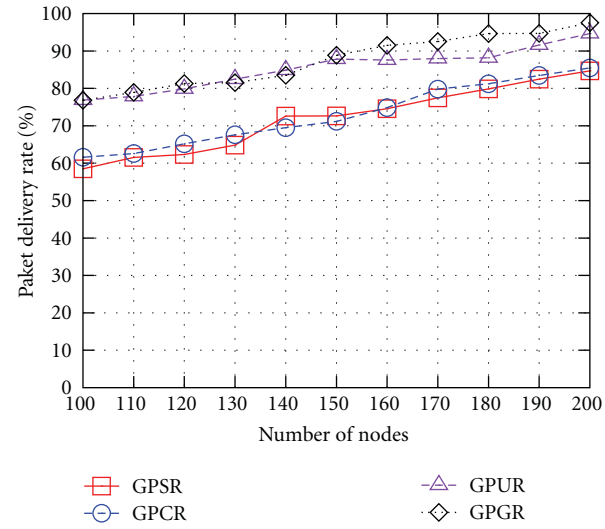


Figure 10: Packet delivery rate based on the number of nodes.

positions of relay candidates in grid sequences of roads, and it estimates the movement directions of nodes.

Figures 8 and 9 show the link breakage rate depending on the number of nodes and the link breakage rate depending on the velocity variation of the nodes, respectively. Figure 8 shows that a larger number of nodes led to lower link breakage rates in all of the routing algorithms. However, the link breakage rate was lower with GPGR compared with GPSR, GPCR, and GPUR. This is because GPGR selects the relay node based on the grid sequence of roads and the movement direction of nodes. With GPSR, GPCR, and GPUR, link breakage occurred because they selected stale nodes as relay nodes that were outside the transmission range, which is a common problem in greedy forwarding. Figure 9 shows that a higher velocity of nodes results in a higher link breakage with all algorithms except GPGR. The link breakage rate when using GPGR was almost constant compared with GPSR, GPCR, and GPUR. This is because GPGR more accurately predicts the position of nodes when selecting the relay node on the grid sequences, even if the velocity of nodes is increased. Thus, there was a lower rate of link breakage due to velocity changes compared with other routing algorithms. With GPSR, GPCR, and GPUR, link breakage occurred with rapid changes in the movement direction of nodes, depending on the increase in node velocity.

Figures 10 and 11 show the packet delivery rate depending on the number of nodes and the packet delivery rate depending on the velocity variation of the nodes, respectively. The packet delivery rate with GPGR was higher compared with GPSR, GPCR, and GPUR. Figure 10 shows that the performance of GPGR was similar to that of GPUR. However, GPGR had better performance than GPUR with an increasing number of nodes. This was because GPUR only

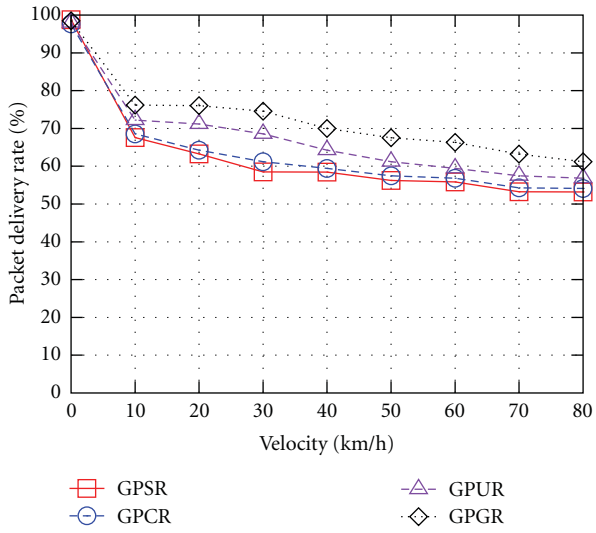


Figure 11: Packet delivery rate based on the velocity of nodes.

considers the location of 2-hop nodes when selecting relay nodes, and it does not consider the road topology. However, GPGR provides a high delivery rate when selecting relay nodes based on the grid sequences of roads. Figure 11 shows that all protocols exhibited falls in the packet delivery rate with increasing node speed. However, GPGR had a lower rate of decrease than the other protocols. This is because GPGR knows the exact position of nodes, so it is expected that the delay time when determining the position of nodes will be lower than other protocols.

5. Conclusion

This paper presented a grid-based predictive geographical routing (GPGR) protocol for urban VANETs. GPGR is an inter-vehicle routing protocol that is based on mobility information from vehicles and a digital urban map, which improves the performance of IVC in VANETs. The GPGR protocol reduces the possibility of link breakage and a local maximum by selecting relay nodes based on mobility information and road topology. Simulation results showed that GPGR produces a very low likelihood of a local maximum, low link break probabilities, and a high packet delivery rate compared with GPSR, GPCR, and GPUR for VANETs. In the future, we will consider the probability of a local maximum and the packet breakage rate in simulations. We will also incorporate more realistic factors into our routing protocol, such as the density of vehicles and delay tolerant characteristics.

References

[1] Q. Tse, "Improving message reception in VANETs," in *Proceedings of the International Conference on Mobile Systems, Applications, and Services (MobiSys '09)*, p. ACM Press, Kraków, Poland, June 2009.

[2] M.-H. Chen, C.-R. Dow, S.-C. Chen, Y.-S. Lee, and S.-F. Hwang, "HarpiaGrid: a geography-aware grid-based routing protocol for vehicular Ad Hoc networks," *Journal of Information Science and Engineering*, vol. 26, no. 3, pp. 817–832, 2010.

[3] J. Luo, X. Gu, T. Zhao, and W. Yan, "A mobile infrastructure based VANET routing protocol in the urban environment," in *Proceedings of the International Conference on Communications and Mobile Computing (CMC '10)*, pp. 432–437, Shenzhen, China, April 2010.

[4] 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.

[5] C. Lochert, M. Mauve, H. Fler, and H. Hartenstein, "Geographic routing in city scenarios," *SIGMOBILE Mobile Computing and Communications Review*, vol. 9, no. 1, pp. 69–72, 2005.

[6] M.-W. Ryu, S.-H. Cha, and K.-H. Cho, "A vehicle communication routing algorithm considering road characteristics and 2-hop neighbors in urban areas," *Journal of Korea Information and Communications Society*, vol. 36, no. 5, pp. 464–470, 2011.

[7] IETF RFC3561, "Ad hoc on-demand distance vector routing," http://www.faqs.org/rfcs/rfc3561.html.

[8] IETF RFC 4728, "Dynamic source routing protocol for mobile ad hoc network for IPv4," http://www.faqs.org/rfcs/rfc4728.html.

[9] J. Nzouonta, N. Rajgure, G. Wang, and C. Borcea, "VANET routing on city roads using real-time vehicular traffic information," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 7, pp. 3609–3626, 2009.

[10] C. Lochert, H. Hartenstein, J. Tian, H. Fler, D. Herrmann, and M. Mauve, "A routing strategy for vehicular ad hoc networks in city environments," in *Proceedings of the Intelligent Vehicles Symposium*, pp. 156–161, IEEE Press, 2003.

[11] W. Sun, H. Yamaguchi, K. Yukimasa, and S. Kusumoto, "GVGrid: a QoS routing protocol for vehicular ad hoc networks," in *Proceedings of the 14th International Workshop on Quality of Service (IWQoS '06)*, pp. 130–139, IEEE Press, New Haven, Conn, USA, June 2006.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 681273, 6 pages
doi:10.1155/2012/681273

Research Article

Reliable Latency-Aware Routing for Clustered WSNs

Ali Tufail

Graduate School of Information and Communication, Ajou University, Suwon 443-749, Republic of Korea

Correspondence should be addressed to Ali Tufail, ally.tufail@gmail.com

Received 14 September 2011; Accepted 2 January 2012

Academic Editor: Tai Hoon Kim

Wireless sensor networks (WSNs) are persistently evolving from merely a notion of microelectronics to a new realm of practical applications. Certain critical applications like disaster management, healthcare, and military not only require exceptionally reliable but also a low-latency source to sink communication. Nevertheless, low source to sink latency is of utmost importance in these kinds of applications. In this paper, a unique latency-sensitive reliable routing protocol for WSNs has been proposed. This protocol uses the concept of hotlines (highly reliable links) and also utilizes alternative routes to reduce the number of hops from the source to the sink. This reduction of hops not only reduces the end-to-end latency but also increases the reliability manifold. The proposed protocol is evaluated with the help of simulation. The simulation suggests that the proposed routing protocol outperforms previously suggested routing protocols in terms of average end-to-end latency and reliability.

1. Introduction

Wireless Sensor Networks (WSNs) have gained focus in last decade or so. The reason of this increased popularity is the support for wide array of crucial application including military, healthcare, industrial monitoring, target tracking, smart homes, and habitat monitoring [1, 2]. A WSN is typically composed of tens to thousands of resource-constrained sensor nodes. These sensor nodes are usually deployed to observe a target phenomenon and report it back to the sink. Besides other resource constraints these nodes have limited battery, signaling, and communication capacity. All these limitations make successful source to sink communication all the more exigent [3].

WSNs and its unique characteristics require special protocols to be developed which are different than those of the traditional wireless or ad hoc networks. Reliable source to sink communication is of the top concern in WSNs [4]. Reliable communication can be very crucial for applications like health monitoring where any failure in the communication can result in a big loss [5]. In a traditional wireless or wired network the communication and the source to the destination link are usually less prone to failure or interference. On the contrary, in WSNs the multihop communication is dependent on the intermediary nodes and any failure in the node or the link can result in failure of the communication. Apart from reliability, latency can be another important factor to be considered in WSNs. Lower source to sink latency can serve as the backbone of the crucial life saving applications.

This paper presents a reliable and latency aware-routing protocol. The proposed protocol works in the clustered WSN environment. Each cluster is managed and controlled by a resource-rich gateway node (GN). In order to have a reliable source to sink communication the proposed protocol makes use of the hotlines that exist between the GNs [6]. Although GN centric routing provides reliability, in certain cases it adds additional latency. This latency is more obvious in cases where the source and the sink lie around the border of the clusters. In order to make this protocol reliable and latency aware this paper presents a border node (BN) routing alternative. The BN routing will not only reduce the source to sink latency but would also increase the reliability by decreasing the number of wireless hops.

The rest of the paper is organized as follows. Section 2 describes related work. Section 3 discusess our network model and assumptions. Section 4 introduces the proposed reliable and low-latency routing protocol. Section 5 gives an overview of the simulation results. Section 6 summarizes key conclusions of this work.

2. Related Work

While reliability of homogeneous WSNs has been investigated in prior studies [7–10], reliability analysis of backbone approach in WSN is largely unexplored. In [8], authors formulate a WSN reliability measure that considers the aggregate flow of sensor data into a sink node. This measure is based on a given estimation of the data generation rate and the failure probability of each sensor. Common-cause failures (CCFs) have been discussed and identified as the cause of unreliability in WSN in [9, 10]. Authors in [9] consider the problem of modeling and evaluating the coverage-oriented reliability of WSN subject to CCF whereas in [10] the main emphasis is on addressing the problem of modeling and evaluating the infrastructure communication reliability of WSN. In [7], the authors compute a measure for the expected and maximum message delay between data sources and data sinks in an operational distributed sensor network (DSN). Authors in [11] present a reliable routing protocol that forms the reliable routing path by utilizing network topology and routing information of the network. However, their protocol and analysis are application specific. Moreover, they have not provided any comparison with the existing reliability protocols.

We can find quite a few papers in the literature that deal with the issue of latency issue of WSNs but to the best of my knowledge this is the first paper that directly addresses the issues of reliability together with the latency in the domain of WSNs. In [12] authors discuss the latency issue in conjunction with the power-saving mechanism. Authors in [13] try to solve the data aggregation issue in WSNs by providing an energy-efficient protocol and claim that the protocol provides reduced latency.

In [14] we presented a Signature-Based Routing protocol (SBR). The purpose of that protocol was also to reduce the end-to-end latency. However, SBR is heavily dependent on the GN and puts much more computation and monitoring burden at the GN. Moreover, in SBR the source node first has to establish a path with the default GN irrespective of its position near the BN. It is only after the GN observes latency much higher than the average latency that the source node is suggested to form an alternative route via BN. This paper introduces BN advertisement messages and BN association mechanism. This new mechanism helps reduce end-to-end latency even further. There is no need to establish a path with the default GN first. Additionally, the extra computation and monitoring overhead has been taken off the GN.

3. Network Architecture

This paper considers a WSN with two-level heterogeneity. The first level has resource-constrained sensor nodes which are deployed densely on a two-dimensional grid. All of the first level nodes have the same resources. The second level has sensor gateway nodes (GNs) that operate as cluster heads to regulate the flow of traffic and to manage sensor motes deployed in the given geographical region. The GNs are not resource constrained and their density is many orders of magnitude lesser than the density of the sensor nodes. GNs are connected to each other in a bus topology via highly reliable links (hotlines), for example, Ethernet cables or point-to-point wireless links. The network topology is assumed to be fixed where both GNs and the sensor nodes are static. The communication from source to destination is multihop. Every node in the network is associated with at least one of the GNs. All the communication, by default, to and from the cluster is routed through the GN, with the exception of the border nodes (BNs). Each BN can connect two clusters by providing an alternative route for intercluster communication. As mentioned in Section 4, BNs would provide lesser latency and at the same time would provide more reliability by reducing the number of wireless hops from the source to the destination.

4. Overview of the Proposed Protocol

4.1. Clustering. Clustering is usually done to serve a particular objective like scalability, load balancing, latency, network management, and so forth [15, 16]. Our suggested approach is tailored to work on the clustered WSNs. Clusters are formed using the same technique as explained in [6]. Each cluster is managed by a gateway node (GN). GN acts as the cluster head for its cluster. All the nodes in a WSN must associate with any one of the GNs and set it as its default. It is worthwhile to mention here that nodes decide to join a particular GN depending upon the hop count. Hop count is calculated by every node after receiving the Router Advertisement (RA) message by the GN. In other words we can say that the closest of all the gateways will have a high probability to serve as the default GN for a particular node.

4.2. Border Nodes. Border nodes (BNs) are the nodes that lie in the overlapping region of clusters. These are shown in Figure 1. BNs are different than the other nodes in the cluster because (a) each BN receives the RA message from more than one GN [6] and (b) each BN can connect two clusters by providing an alternative route for intercluster communication. This paper explores the second characteristic of the BNs to provide lesser latency at the same time providing more reliability by reducing the number of wireless hops from the source to the destination.

There are two ways to communicate between a source and a destination. The default is the source-GN-GN-destination route. The other possible route is source-BN-destination. Later alternative route is expected to be suitable only for the first hop nodes of the BNs; otherwise it might yield more latency and less reliability. The route options are shown in Figure 3. On the other hand Figure 2 highlights the border nodes and its first hop neighbors on both sides.

4.3. Border Node Association. In order to have a communication using BNs, there must be some association mechanism that binds the first hop neighbors of the BNs to the BNs. This paper defines BN association mechanism that is similar to the GN association mechanism [6] except (a) BN association mechanism which is just targeted to the first hop neighbors

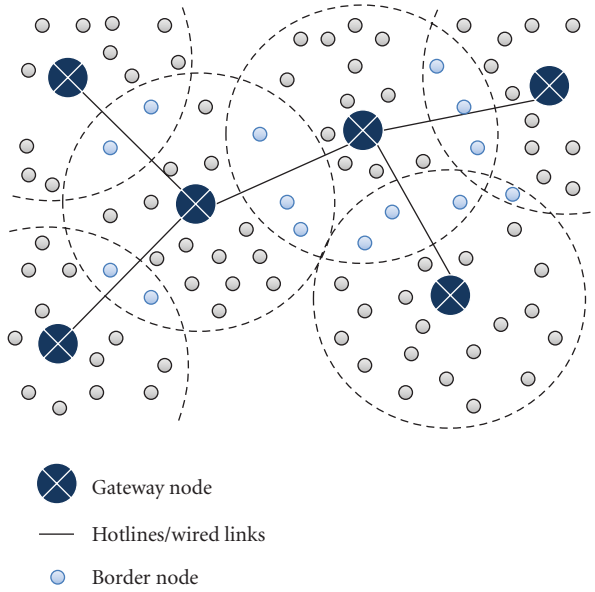


Figure 1: Network architecture.

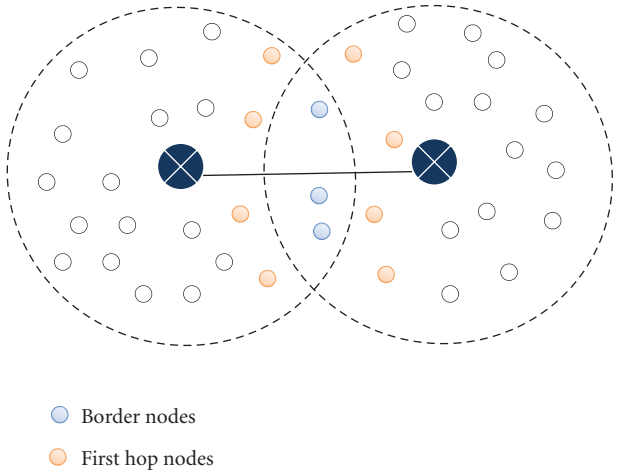


Figure 2: Border nodes.

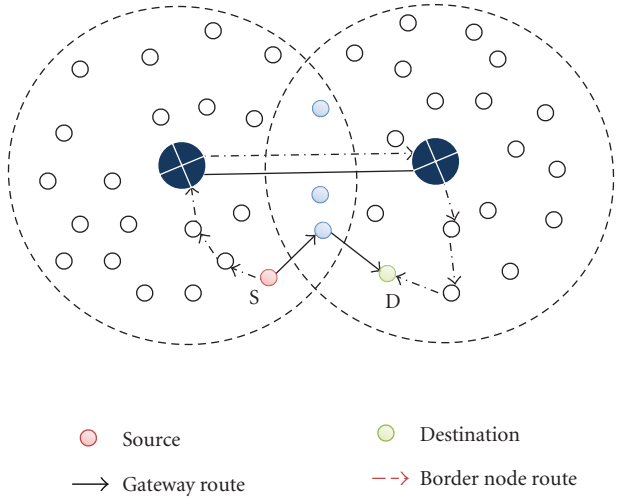


Figure 3: Multiple routes from the source to the destination.

of the BNs and (b) BN advertisement message (BNAs) is much less frequent as compared to GN's RA message.

First hop neighbor nodes can lie in the communication range of more than one BNs, but they associate to only one BN. First hop neighbor node decides its association with the help of first BNA it receives from a particular BN. First hop neighbor node keeps the other BNs information in its routing table so that if later the default BN is not accessible it still can communicate with the destination using the backup route.

Please note that there are two different associations within one cluster. One is for the ordinary sensor nodes. They get associated to the GN. If these nodes want to communicate they have to communicate using the GN. Other association is for the first hop neighbors of the BNs. The first hop neighbors have in a sense two associations, one with the GN and other with the next hop BN. For these kinds of nodes the communication from source to the destination can go via either GN or BN. Here it should be clear that BNs are also associated with at least one of the GNs. Depending upon the destination BNs can communicate either with the help of the GN or directly with the destination.

4.4. End-to-End Communication. If a source has to communicate with a destination, by default, it will discover the route towards the GN. This will not be true for the first hop neighbors of the BNs. For these nodes, BNs are fist consulted. If the destination is one hop away from the BN it will tell the source to use BN as its default route, otherwise BN will reply in negative and the source has to use GN route to reach the destination. If the BN route is being used it means the source and destination are just two hops away. If we compare this route with the default GN route, we can see that in this case BN route is providing less latency and more reliability. Reliability is increased because end-to-end number of wireless hops is reduced.

Algorithm 1 explains how a source node tries to find out the route to the destination node.

5. Simulation and Results

This section discusses the simulations and their results. Table 1 explains simulation setup and the environment that has been utilized to conduct the simulations.

The simulations have been conducted by having the same network topology as defined in Section 3. All the traffic goes to the default GN except for the first hop neighbor of the BNs. The first hop neighbor of the BN first asks the BN about the destination if BN has the destination in its reach (within its first hop) then the route is established using BN; otherwise default GN route is established. All the GNs are connected via Ethernet cable. Multiple sources send data to the sink to emulate bursty traffic.

```
Legends
BN: Border Node, GN: Gateway Node, A: Source Node, B: Destination Node
(1)  Start Communication—find a path from A to B
(2)    If A is first hop neighbor of the BN
(3)      Find B
(4)        If B is first hop neighbor of the BN
(5)          Establish a path from A to B using BN
(6)        Else reach the default gateway
(7)          Establish a path from A to B using A's and B's default GNs
(8)    Else
(9)      Find B
(10)       Establish a path using default GN
```

Algorithm 1: Algorithm to find a path from the source to the destination.

Table 1: Simulation environment.

Simulation environment	Qualnet 4.5
Routing protocol	AODV
Intergateway routing	OSPF
Intracluster communication	Wireless (802.15.4)
Intercluster communication	Wired (Ethernet)
Number of nodes	100
Total terrain area	1500 m × 1500 m
Simulation time	4200 seconds
Total runs	30

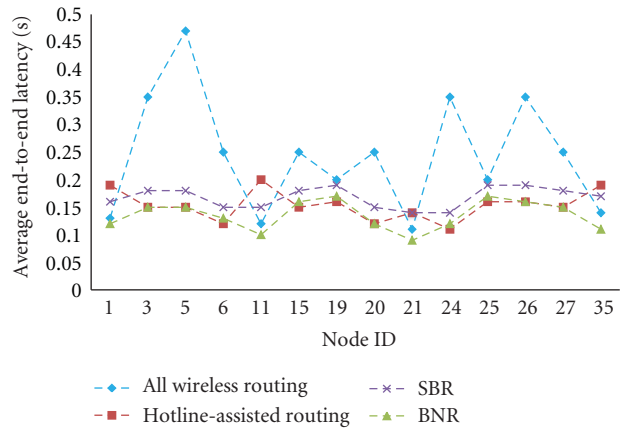


Figure 4: Comparison of average end-to-end delay.

For simplification and comparison purpose in this section the suggested scheme is called as Border Node assisted Routing (BNR). BNR is compared with the all wireless routing, hotline-assisted routing [6], and signature-based routing (SBR) [14]. For all wireless routing same network topology and node deployment have been kept. However, there are no GNs or Ethernet connections. For hotline-assisted routing GNs are deployed along with their Ethernet connections. For SBR latency-aware routing is utilized along with the hotline-assisted routing. Please note that the routing used in BNR is also hotline assisted but there is a difference in finding and establishing the source to destination route.

This section defines reliability as the percentage of total messages received successfully by the destination. Latency is defined as the time taken by a packet to reach from the source to the destination/sink. For better comparison it also includes time consumed to establish or to change/reestablish (in case of SBR) the route from source to destination.

Figure 4 compares all the protocols in terms of average end-to-end latency against node IDs. Please note that the node IDs are the same as in our experiment, only the nodes of one cluster are shown here. In Figure 4 nodes 11, 21, and 35 are the first hop neighbors of the BNs. We can see that all wireless routing is giving highest average latency for all the nodes except for the nodes 11, 21, and 35. The reason for this high latency for all the other nodes is that the traditional way of routing just includes wireless hops. Each wireless hop increases a certain amount of latency. However, nodes 11, 21, and 35 happen to have their destination quite close by, therefore they provide less latency as compared to the hotline approach. On the contrary, hotline-assisted routing is better than the traditional all wireless routing for most of the cases as it provides high-speed alternative to the wireless hops and reduces the total number of end-to-end hops. SBR and BNR both perform better than the basic hotline approach in cases where the source node is the first hop neighbor of the BNs; otherwise the performance is almost similar to the basic hotline approach. BNR outperforms SBR because the path establishment is simpler in case of BNR and path re-establishment phase has been completely removed from BNR which decrease the latency even further.

Figure 5 compares the protocols in changing network density. The figure summarizes the average of all the communication that took place during all the simulation runs for each protocol. We can see that all the three protocols are better than the traditional wireless routing. However, they are close to each other in terms of performance. BNR is slightly better than hotline-assisted routing and SBR due to the already explained fact that it considers an alternative routing path for the nodes which are first hop neighbors of the BNs. Although SBR also considers alternative routing path, the process to find that path is complicated and takes more time as compared to BNR.

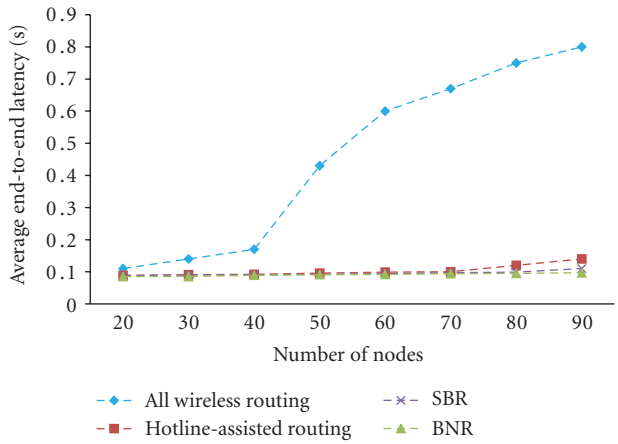


Figure 5: Average end-to-end latency with increasing network density.

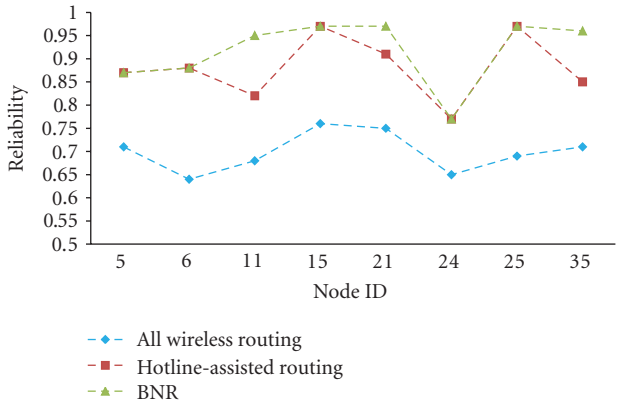


Figure 6: Comparison of reliability.

Figure 6 compares the reliability for all the protocols. SBR and BNR have similar reliability as they have almost similar path from source to the destination. We can see that BNR is better for the first hop neighbor nodes of the BNs as it provides a shorter and more reliable path. With the increase of a wireless hop the probability of losing a packet increases so BNR is giving more reliability for nodes 11, 21, and 35. For rest of the nodes it provides same reliability as hotline-assisted routing but it is much better than the traditional wireless routing.

6. Conclusion

This paper talks about reliability and latency issues of WSNs. Reliability has been improved with the help of hotlines between the gateway nodes (GNs). However, in order to make that protocol latency aware or in other words in order to decrease end-to-end latency, a latency-aware routing protocol has been introduced. This protocol does not only decrease the latency from source to the destination but also enhances the reliability by reducing number of end-to-end wireless hops. The suggested protocol introduces alternative path from the source to the destination by utilizing the border nodes (BNs). BNs also take the computation and monitoring overhead off by serving as the alternative forwarder or communicator between two clusters. The simulation results also prove that the suggested routing approach outperforms the previously suggested protocols by decreasing the latency and increasing the packet success ratio.

Acknowledgment

This work was supported by the new faculty research fund of Ajou University, Suwon, South Korea.

References

[1] F. Fiorenzo, G. Maurizio, M. Domenico, and M. Luca, "A review of localization algorithms for distributed wireless sensor networks in manufacturing," *International Journal of Computer Integrated Manufacturing*, vol. 22, pp. 698–716, 2009.

[2] Y. C. Wang, Y. Y. Hsieh, and Y. C. Tseng, "Multiresolution spatial and temporal coding in a wireless sensor network for long-term monitoring applications," *IEEE Transactions on Computers*, vol. 58, no. 6, pp. 827–828, 2009.

[3] T. T. Long and J. Levendovszky, "A novel reliability based routing protocol for power aware communications in wireless sensor networks," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '09)*, pp. 1–6, Budapest, Hungary, April 2009.

[4] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–105, 2002.

[5] U. Varshney and S. Sneha, "Patient monitoring using ad hoc wireless networks: reliability and power management," *IEEE Communications Magazine*, vol. 44, no. 4, pp. 49–55, 2006.

[6] A. Tufail, S. A. Khayam, S. D. Hwan, and K. H. Kim, "A hotline-based reliable topology for wireless sensor networks," in *Proceedings of the 3rd International Conference on Sensor Technologies and Applications (SENSORCOMM '09)*, pp. 562–567, Athens, Greece, June 2009.

[7] H. M. F. AboElFotoh, S. S. Iyengar, and K. Chakrabarty, "Computing reliability and message delay for cooperative wireless distributed sensor networks subject to random failures," *IEEE Transactions on Reliability*, vol. 54, no. 1, pp. 145–155, 2005.

[8] H. M. F. AboElFotoh, E. S. ElMallah, and H. S. Hassanein, "On the reliability of wireless sensor networks," in *Proceedings of the IEEE International Conference on Communications (ICC '06)*, pp. 3455–3460, Istanbul, Turkey, July 2006.

[9] A. Shrestha, X. Liudong, and L. Hong, "Modeling and evaluating the reliability of wireless sensor networks," in *Proceedings of the 53rd Annual Reliability and Maintainability Sympsoium (RAMS '07)*, pp. 186–191, Orlando, Fla, USA, January 2007.

[10] A. Shrestha, X. Liudong, and L. Hong, "Infrastructure communication reliability of wireless sensor networks," in *Proceedings of the 2nd IEEE International Symposium on Dependable, Autonomic and Secure Computing (DASC '06)*, pp. 250–257, Indianapolis, Ind, USA, October 2006.

[11] J. Dong, W. Qianping, Z. Yan, and W. Ke, "The research and design of high reliability routing protocol of wireless sensor network in coal mine," in *Proceedings of the International Conference on Networks Security, Wireless Communications and*

Trusted Computing (NSWCTC '09), pp. 568–571, Wuhan, China, April 2009.

[12] O. Dousse, P. Mannersalo, and P. Thiran, "Latency of wireless sensor networks with uncoordinated power saving mechanisms," in *Proceedings of the 5th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MoBiHoc '04)*, pp. 109–120, May 2004.

[13] Y. Li, L. Guo, and S. K. Prasad, "An energy-efficient distributed algorithm for minimum-latency aggregation scheduling in wireless sensor networks," in *Proceedings of the 30th IEEE International Conference on Distributed Computing Systems (ICDCS '10)*, pp. 827–836, Genova, Italy, June 2010.

[14] A. Tufail, S. A. Khayam, M. T. Raza, A. Ali, and K. H. Kim, "An enhanced backbone-assisted reliable framework for wireless sensor networks," *Sensors*, vol. 10, no. 3, pp. 1619–1651, 2010.

[15] W. Youssef and M. Younis, "Intelligent gateways placement for reduced data latency in wireless sensor networks," in *Proceedings of the IEEE International Conference on Communications (ICC '07)*, pp. 3805–3814, Glasgow, UK, June 2007.

[16] C. P. Low, C. Fang, J. Mee, and W. Hock, "Load-balanced clustering algorithms for wireless sensor networks," in *Proceedings of the IEEE International Conference on Communications (ICC '07)*, pp. 3485–3490, Glasgow, UK, June 2007.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 790131, 10 pages
doi:10.1155/2012/790131

Research Article

A Dynamic Traffic-Aware Duty Cycle Adjustment MAC Protocol for Energy Conserving in Wireless Sensor Networks

Tz-Heng Hsu,[1] Tai-Hoon Kim,[2] Chao-Chun Chen,[1,3] and Jyun-Sian Wu[1]

[1] *Department of Computer Science and Information Engineering, Southern Taiwan University, Tainan, Taiwan*
[2] *GVSA and University of Tasmania, Tasmania, Australia*
[3] *Institute of Manufacturing Information and Systems, National Cheng Kung University, Tainan, Taiwan*

Correspondence should be addressed to Tai-Hoon Kim, taihoonn@paran.com

Received 27 October 2011; Accepted 9 November 2011

Academic Editor: Sabah Mohammed

Wireless sensors are battery-limited sensing and computing devices. How to prolong the lifetime of wireless sensors becomes an important issue. In order to reduce the energy consumptions when nodes are in idle listening, duty-cycle-based MAC protocols are introduced to let node go into sleep mode periodically or aperiodically. The long duty cycle makes sensors increase the transmission throughput but consumes more energy. The short duty cycle makes sensors have low energy consumption rate but increases the transmission delay. In this paper, a dynamic traffic-aware MAC protocol for energy conserving in wireless sensor networks is proposed. The proposed MAC protocol can provide better data transmission rate when sensors are with high traffic loading. On the other hand, the proposed MAC protocol can save energy when sensors are with low traffic loading. Simulation results show that the proposed protocol has better data throughput than other duty-cycle-based MAC protocols, for example, S-MAC and U-MAC. We also developed a set of comprehensive experiments based on the well-known OMNET++ simulator and revealed that our proposed TA-MAC performs significantly outstanding than related schemes under various situations.

1. Introduction

A wireless sensor network (WSN) is composed of hundreds or thousands of autonomous and compact sensing devices, gathering and delivering information about events of interest to observers. In a wireless sensor network, the energy efficiency of communications among sensors is a major issue. Various energy conservation schemes for sensor networks have been proposed [1–5]. TDMA-based protocols, for example, LEACH [1] and BMA [2], provide contention-free mechanisms for sensors to avoid the competition of the radio channels. The major advantage of TDMA-based protocols is energy efficiency due to no energy wasted on collisions caused by channel contention [6–8]. However, classic TDMA protocols are hard against the unpredictable variations of sensor networks.

In traditional wireless sensor networks, sensors need to be in listening for sensing data transmission. If no sensing event happens, nodes are in idle for a long time. Idle listening wastes energy when a node is active, even there is no meaningful task on the radio channel. Traditional MAC protocols, such as 802.11, are unsuitable for sensor networks' data delivery. Idle listening in 802.11 consumes as much energy as it does when receiving data. In order to reduce the energy consumption when nodes are in idle listening, duty-cycle-based MAC protocols are introduced to let nodes go into sleep mode periodically or aperiodically [9–13].

S-MAC is a MAC protocol that can specify when nodes are awake and asleep within a frame [9]. In S-MAC, the period of a regular cycle consists of a listen and a sleep state. The duty cycle is defined as the ratio of the node in listen state compared with the period of a regular cycle. Within a virtual cluster, a node exchanges synchronization and schedule information with its neighbors to ensure that the node and its neighbors wake up at the same time. In S-MAC, a uniform duty cycle is assigned across the whole network. The uniform duty cycle assignment is not suitable for all sensors because not all sensors have the same duty or workloads. If the uniform duty cycle is set too long, it may cause energy wastage on nodes with low data traffic because the nodes are in idle listening. If the uniform duty cycle is set too short, it may increase transmission latency on nodes

with heavy data traffic because they do not have enough time to transmit all collected data [9]. T-MAC protocol enhances the design of S-MAC by listening to the radio channel for only a short time, and then nodes enter into sleep mode if no data is received during that time [10]. After synchronization, T-MAC shortens the awake period of a node if the radio channel is idle for a short time. If data is received, the node remains awake until the transmission is finished or the awake period ends. By using adaptive duty cycling, T-MAC can reduce energy usage for data transmission. However, in return for gaining energy, the costs are the low data throughput and the long transmission latency.

U-MAC protocol tunes duty cycle according to the traffic loads of sensor nodes. In wireless sensor networks, making nodes asleep for a long time to lower the energy consumption may result in higher transmission latency [11]. U-MAC tries to solve the problem by tuning duty cycle according to sensors' traffic loads. To reflect the traffic loads, U-MAC adopts a utilization function to calculate the load of each node. Upon synchronization, a node calculates its traffic utilization since the last synchronization time. The node adjusts its duty cycle according to the calculated utilization and then broadcasts the new schedule to its neighbors. The performance evaluation shows that the U-MAC can have good throughput when traffic loading is high. Meanwhile, U-MAC can save more energy than S-MAC when traffic loading is low. However, the problems of duty cycle synchronization between sender and receiver are not well solved in U-MAC. For a forwarding sensor, the duty cycle will be increased for getting more time to transfer data. For a receiving sensor with low traffic utilization, the duty cycle will be decreased. After a long time, the forwarding sensor is working in the high duty cycle and the receiving sensor may work in the low duty cycle. When the forwarding sensor sends a packet to a receiving sensor, the packet may be lost due to the receiving sensor enters into its sleep period early. As a result, channel utilization and transmission delay are varying in U-MAC.

In this paper, we propose a dynamic traffic-aware duty cycle adjustment MAC protocol named "TA-MAC" which can adjust sensor's duty cycle adaptively according to status of sending/receiving buffer, traffic loading, and battery life. After the information of a sensor is obtained, the sensor adaptively adjusts the duty cycle for sending and receiving packets. The goals of the proposed MAC protocol are twofold: (i) to conserve energy on sensors with low traffic loading and (ii) to decrease transmission latency and increase data throughput on sensors with heavy traffic loading. Each sensor can have different duty cycle for data transmission, long duty cycle is allocated for sensors that have more data need to be sent in sending queue. Each sensor extends its duty cycle if the sensor has data need to transmit. For sensors that failed to send data due to radio channel competition, these sensors will go to sleep state for energy conserving and wait for next wake-up time to compete the radio channel again for transmitting data.

The rest of the paper is organized as follows. Section 2 describes the related works. Section 3 presents the proposed dynamic duty cycle adjustment MAC protocol. Section 4 shows the experimental results. Section 5 concludes the proposed MAC scheme.

2. Related Works

This section provides a brief overview of existing energy conserving MAC protocols in asynchronous and synchronized approaches.

For asynchronous protocols, for example, B-MAC [12], WiseMAC [13], and X-MAC [14], they are utilizing adaptive preamble sampling schemes to reduce duty cycle and minimize idle listening. In B-MAC, nodes are awake and asleep asynchronously. When a sender has data, the sender sends a packet with a preamble longer than the sleep period of the receiver. When the receiver wakes up and detects the preamble, it stays awake to receive the data. B-MAC equips periodic channel sampling, called low-power listening (LPL), to enable low-power communication, without the need of explicit synchronization among the nodes. The receiver only wakes for a short time to sample the medium, which reduces the energy consumptions. In [13], El-Hoiydi and Decotignie proposed the WiseMAC protocol for the downlink channel of infrastructure wireless sensor networks. WiseMAC is based on the preamble sampling technique [15]. By sampling the medium, WiseMAC listens to the radio channel for a short duration, for example, the duration of a modulation symbol, to check for network activities. Nodes sample the medium with the same constant period. In WiseMAC, the access point learns the sampling schedule of all sensor nodes. Having the sampling schedule of the destination, the access point starts the transmission with a wake-up preamble. To reduce the length of the extended preamble, the receiver sends the time of its next awake period in the data acknowledgment frame. By using the preamble sampling technique, WiseMAC can reduce energy usage for transmitting data. However, WiseMAC is not suitable for long distance multihop data transmission. In [14], Buettner et al. proposed the X-MAC protocol that uses a shortened preamble approach for retaining the benefits of low-power listening. In wireless sensor networks, long preamble introduces excess latency at each hop and suffers from excess energy consumption at nontarget receivers. Instead of sending a long preamble stream, X-MAC inserts small pauses in the preamble stream, which forms a series of short preamble streams. These gaps enable the receiver to reply an early acknowledgement packet back to the sender during the short pause between preamble streams. When a sender receives an acknowledgement packet from the receiver, it stops sending preamble and then starts the transmission of data packet. Nontarget receivers who overhear the strobed preamble can go back to sleep immediately. X-MAC's shortened preamble approach can reduce energy usage at both the transmitter and receiver side. X-MAC also proposed an algorithm for adapting the duty cycle of the receiver to adapt to different traffic loads. However, the effect of the proposed adaptive algorithm is not clear in the performance evaluation.

Synchronized protocols, such as S-MAC [9], T-MAC [10], and U-MAC [11], can specify when nodes are awake

and asleep within a frame. In S-MAC, neighboring nodes form virtual clusters to synchronize the working schedules [7]. Nodes periodically wake up, receive and transmit data, and then return to sleep to reduce energy consumptions. Within a virtual cluster, a node exchanges synchronization and schedule information with its neighbors to ensure that the node and its neighbors wake up at the same time. After the schedule is synchronized, nodes with data can transmit packets using RTS (Request To Send) and CTS (Clear To Send) packets when the nodes awake. Nodes without data will be in idle state for waiting connection until the end of the awake period and the nodes then enter sleep mode. S-MAC uses a SYNC packet to accomplish the synchronization of working schedule. When a node receives the SYNC packet, the node will adjust its timer immediately. In S-MAC, the listen interval is divided into two parts: (i) interval for receiving SYNC packets and (ii) interval for receiving RTS/CTS packets.

In [10], Van Dam and Langendoen proposed the T-MAC protocol that enhances the design of S-MAC by listening to the channel for only a short time, and if no data is received during this window, the node enters sleep mode. After the synchronization, T-MAC shortens the awake period of a node if the channel is idle for a short time. Once a data is received, the node remains awake until the transmission is finished or the awake period ends. By using adaptive duty cycling, T-MAC can reduce energy usage for data transmission. As a tradeoff, the costs are the reduced data throughput and the increased transmission latency. In [11], Yang et al. proposed the U-MAC protocol that tunes duty cycle according to the traffic loads of sensors. In a wireless sensor network, making nodes asleep for a long time to lower the energy consumption may result in higher latency. U-MAC tries to solve the problem by tuning sensor's duty cycle according to the traffic loads and selective sleeping after transmission. To reflect the traffic loads, U-MAC adopts a utilization function to calculate the loads of each sensor. Upon synchronization, the sensor calculates its traffic utilization since the last synchronization time. The sensor adjusts its duty cycle according to the calculated utilization and then broadcasts the new schedule to its neighbors.

Most of the recent works focus on how to prolong a sensor's or network's lifetime. However, these works do not consider the performance of data transmission when sensors have different workload. Our work differs from all of the above works by striking the balance between the battery life and transmission latency according to the traffic situations. The main difference is that the proposed MAC protocol is based on the demand of data transmission of sensors. The use of traffic-aware protocol can conserve sensors' battery life and reduce the transmission latency caused by sleep.

3. The Dynamic Traffic-Aware Duty Cycle Adjustment MAC Protocol

Most energy conserving MAC protocols adjust the sensor's duty cycle according to some factors like transmission latency and traffic load. For example, DS-MAC [16] adjusts the sensor's duty cycle according to the data transmission delay of sensors. U-MAC [5] determines the sensor's duty cycle according to sensor's traffic utilization. Some duty cycle adjustment algorithms change the sensor's listening time, others adjust the sensor's frame time [16, 17]. Existing adaptive duty cycle algorithms consider only limited factors to adjust the sensor's duty cycle. The lack of complete information of the network and sensor's sending/receiving status leads to additional energy consumption. Figure 1 shows an example of data transmission in U-MAC with cross topology. In Figure 1, the forwarding node needs to receive data from senders and then transmit the data to receivers. Since U-MAC adjusts the sensor's duty cycle according to the calculations of sensor's traffic utilization, the forwarding node has higher duty cycle than the other nodes. On the other hand, the sending nodes and receiving nodes will have lower duty cycle than the forwarding node. If a receiving node's duty cycle is the same as the forwarding node, increasing duty cycle can indeed improve the data transmission rate. Figure 2 shows that the duty cycle adjustment can enhance the transmission efficiency if the participant sensors increase their duty cycles.

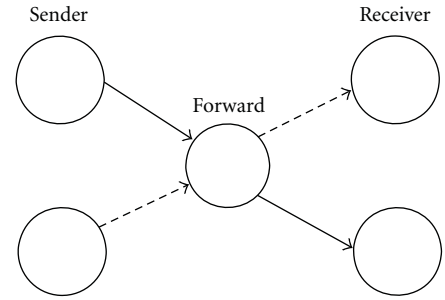


Figure 1: Sensors in U-MAC with cross topology.

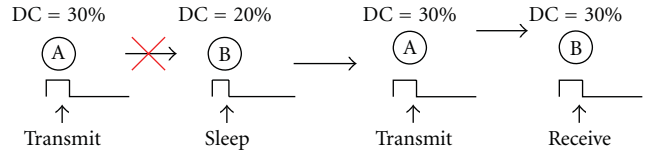


Figure 2: The increasing of sensors' duty cycles raises transmission efficiency.

In U-MAC, senders and receivers try to reduce the duty cycle to save energy consumption if they are in low traffic utilizations. Figure 3(a) shows that all sensor nodes have the initial duty cycle of 20%. Over a long time, Figure 3(b) shows the senders A, B, C and receivers E, F, G reduce the duty cycle because of the low traffic utilizations. However, the forwarding node D increases its duty cycle in high usage status because of node D is communicating with multiple senders and receivers. When the senders and the receivers go to sleep, the forwarding node keeps on trying to transfer data packets from senders to receivers because of the forwarding node is in high duty cycle. In case of senders and receivers in low duty cycle, the forwarding node cannot receive packets from senders because of senders are in sleeping. Meanwhile,

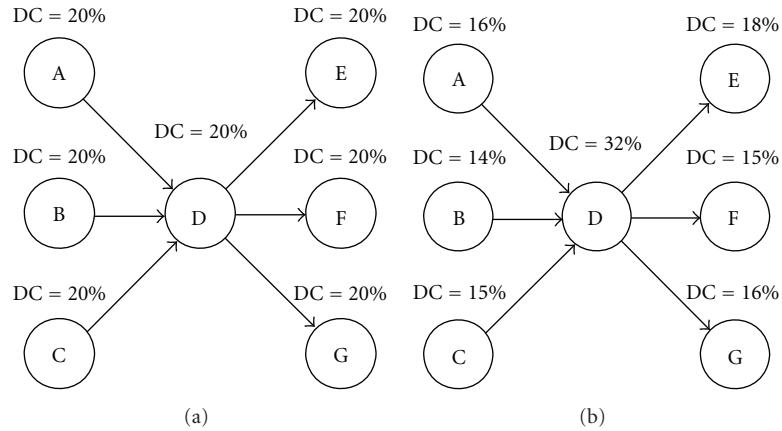


Figure 3: The result of duty cycle adjustment of sensors in U-MAC over a long time.

the forwarding node cannot transfer data packets to receivers because of receivers are in sleeping too. As a result, the forwarding node is wasting energy as it is in idle state.

Reducing the duty cycle of sensors can save energy to prolong sensors' life when sensors are in low traffic loads. Extending the duty cycle of sensors can improve data efficiency and reduce transmission delay when sensors are in high traffic loads. However, the duty cycle synchronization problems are not well solved in U-MAC. As a result, channel utilization and transmission delay are varying case by case in U-MAC.

To overcome the aforementioned issues, we propose a dynamic traffic-aware duty cycle adjustment MAC protocol named "TA-MAC" that can adjust duty cycle adaptively according to status of sensor's sending/receiving buffer, traffic loading, and battery life. The proposed traffic-aware duty cycle adjustment MAC protocol is based on the existing U-MAC protocol, and more functions are designed by us to enhance it. In U-MAC, the forwarding node has higher duty cycle due to its high traffic load. When senders continue to send data packets to receivers, this makes the forwarding node continues to increase the duty cycle to adapt to the current load. However, the senders and receivers do not have such a high traffic load. Therefore, the senders and receivers will lower their duty cycle. To solve such issue, the proposed MAC protocol utilizes an additional control message to achieve duty cycle synchronization between two sensor nodes. The receiving node can decide whether to keep the same duty cycle or not. If the forwarder and receiver have the same duty cycle, the sensors can have more time to transmit data. Based on such idea, the proposed MAC protocol performs outstanding than related schemes under various situations, and the performance comparisons are shown in next section. In the following, we present the technical details of the TA-MAC protocol.

3.1. TA-MAC Algorithm. Algorithms 1 and 2 present the proposed TA-MAC algorithm. Table 1 depicts the symbols used in the proposed algorithm. The initial duty cycle of all nodes is set to the minimum duty cycle value DC_{min}.

Table 1: Symbols used in the TA-MAC algorithm.

Symbol	Meaning
$Queue_{high}$	Queue high
WR	Working rate
WR_{high}	Working rate high
WR_{low}	Working rate low
T_{tx}	Total transmission time
T_{rx}	Total receive time
Tidle	Total idle time
DC	Duty cycle
DC_{max}	The maximum of duty cycle
DC_{min}	The minimal of duty cycle
PL	Power level

The algorithm is divided into two parts: (i) sending duty cycle adjustment packet and (ii) processing the duty cycle adjustment packet.

3.1.1. SendDCAdjust: Sending Duty Cycle Adjustment Packet. When the transfer of a synchronization packet is completed, the node enters into duty cycle adjustment mode, where the adjustments can be further divided into two conditions: (i) normal state and (ii) abnormal state.

The duty cycle adjustment in normal state is based on the sensor's working rate, duty cycle, and the remaining energy. Working rate WR denotes the proportion of working time of a wireless sensor in a period of duty cycle. The working rate WR can be calculated by the total working time divided by the total operating time (active time). If the current working rate WR is higher than the working rate threshold WR_{high}, the current duty cycle is lower than the maximum duty cycle DC_{max}, and the remaining energy is in normallevel, then the duty cycle will be increased by n%. If the current working rate WR is lower than the working rate threshold WR_{low} and the current duty cycle is higher than the minimum duty cycle DC_{min}, then the duty cycle will be decreased by n%. When the traffic load is high, the sender's duty cycle will continue to increase until it reaches the DC_{max} value. The sender will

Symbols definition:
Pkt: the packet that generated by sensor
end definition
1: Initial DC : DC_{min}
2: **while** a SYNC packet is sent **do**
3: Sending_Queue = Sensor.GetSendingQueueLength()
4: **if** Sending_Queue $\geq$ $Queue_{high}$ and Sensor.DC < DC_{max} and Sensor.PL = N and Sensor.State $\neq$ $Queue_{high}$ **then**
5: Sensor.DC = DC_{max}
6: Pkt.Type = DC
7: Pkt.Data = Sensor.DC
8: Send(Pkt)
9: Sensor. State = $Queue_{high}$
10: break
11: **end if**
12: $WR = \frac{T_{tx} + T_{rx}}{T_{tx} + T_{rx} + T_{idle}} * 100\%$
13: **if** Sensor.WR $\geq$ WR_{high} and Sensor.DC < DC_{max} and Sensor.PL = N **then**
14: Sensor.DC = Sensor.DC $*(1 + n\%)$
15: **else if** Sensor.WR $\leq$ WR_{low} and Sensor.DC > DC_{min} **then**
16: Sensor.DC = Sensor.DC $*(1 - n\%)$
17: **end if**
18: Curren_DC = Sensor.DC
19: **if** Currcnt_DC $\geq$ DC_{max} and Sensor. State $\neq$ DC_{max} **then**
20: Pkt.Data = Sensor.DC
21: Send(Pkt)
22: Sensor.State = DC_{max}
23: break
24: **end if**
25: //when the DC is less than DC_{max}, clean the request
26: **if** Sensor.State = DC_{max} and Sensor.DC < DC_{max} **then**
27: Sensor.State = Normal
28: **end if**
29: //when the queue length less than the queue threshold, clean the request
30: **if** Sensor. State = $Queue_{high}$ and Sending_Queue < $Queue_{high}$**then**
31: Sensor.State = Normal
32: **end if**
33: **end while**

Algorithm 1: Send DCAdjust packet.

send a control packet to notify the receiver that the duty cycle is adjusted. The receiver will then adjust its duty cycle if the receiver has enough energy. If the receiver does not have enough energy, the receiver will not change its duty cycle.

The duty cycle adjustment in abnormal state is triggered when some situations occurred, for example, sending/receiving queue is full. Each node has a sending queue and a receiving queue for storing incoming/outgoing packets temporarily. When a packet is generated, the packet is immediately put into the sending queue and waiting for transmission. Since the queue length is limited, when the queue is full, later packets will be dropped due to out of memory, resulting in the reduction of transmission throughput. Figure 4 shows the illustration of packets that are dropped because the sending queue is full.

To overcome the above-mentioned issue, before the sending queue is going to full, the proposed MAC protocol will increase the sensor's duty cycle rapidly to the maximum duty cycle value DC_{max} for prolonging the working time to transmit queued packets. However, the sensor's survival time will be greatly reduced if the sensor's duty cycle is further increasing when a sensor is in low-battery condition. Therefore, the current sensor's remaining energy is an important factor for the duty cycle adjustment. The states of remaining energy can be divided into (i) low-energy state and (2) normal-energy state. As shown in Table 2, L for low level, representatives the remaining energy is currently less than 40% of full battery life. N is normal, representatives the current residual energy is sufficient for the general use. In the case of the sensor is in the low-energy state, the duty cycle cannot be increased in order to prolong the sensor's life time. When the remaining energy is in the normal energy state, the proposed MAC protocol will increase the sensor's duty cycle for sending more queued packets.

Increasing the duty cycle can improve the efficiency of packet transmission; however, the receiver must be synchronized to follow the same schedule in order to receive the data sent from the sender. The duty cycle in the sender and receiver is independent. When the sender's duty cycle is higher than the receiver's duty cycle, it will cause the receiver

```
1: while receiving a packet do
2:     if Pkt.Type = DC then
3:         if SensorDC < DC_max and Sensor.PL = N then
4:             Sensor. DC = Pkt.Data
5:         end if
6:     else
7:         if Pkt.Type = SYNC then
8:             Sync()
9:       end if
10:    else
11:        if Pkt.Type = RTS then
12:            Rts()
13:        end if
14:    else
15:        if Pkt.Type = CTS then
16:            Cts()
17:        end if
18:    else
19:        if Pkt.Type = DATA then
20:            Data()
21:        end if
22:    end if
23: end while
```

Algorithm 2: Handle DCAdjust packet.

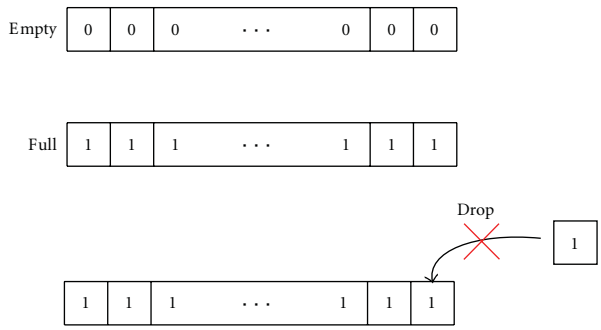


Figure 4: The illustration of packets that are dropped because the sending queue is full.

Table 2: Energy Level.

Name	Low level	High level
Remaining energy	0% ~ 39%	40% ~ 100%
PL	L (low)	N (normal)

to sleep earlier than the sender. As a result, the data cannot be sent to the receiver even the sender is still in working condition. In Figure 5, node A is the sender and node B is the receiver, the packet throughput cannot be increased even the node A increases its duty cycle, because of node B is in sleeping state. Therefore, node A wastes its precious battery power in idle waiting.

For the above problem, the proposed MAC protocol sends an additional control packet to let the sender and receiver have the same duty cycle. In the proposed MAC protocol, a control packet named DCAdjust is defined. In Figure 6, node A and node B can have the same duty cycle after additional duty cycle synchronization. Finally, node A can have more time to transmit data to node B. As a result, data throughput is increased. Figure 7 illustrates the structure of the control packet DCAdjust. The packet is used to tell the receiver about the current state of the sender's duty cycle. The receiver will determine whether to synchronize the sender's duty cycle or not according to its remaining battery power. The structure of the control packet is similar to the IEEE 802.11 RTS (Request To Send) packet. The control packet has a 1-byte field named DC information, which is used to record the sender's duty cycle. Since the length of the control packet DCAdjust is small, the energy spent in sending and receiving the control packet DCAdjust is low.

3.1.2. Handle the DCAdjust Packet. When the receiver is receiving the control packet, DCAdjust, the receiver will adjust its duty cycle according to its remaining energy. If the receiver does not have enough energy, the receiver will not change its duty cycle.

4. Performance Evaluation

In order to evaluate the performance of the proposed control scheme, we use the OMNeT++ discrete event simulator with EYES WSN simulation framework to simulate the wireless sensor networking environment [12]. There are five nodes in a 10 by 10 grid in the simulated sensor model. Nodes exchange messages using the wireless communication. A message will be heard by all the neighbors situated within the transmission range. Three duty-cycle-based MAC protocols

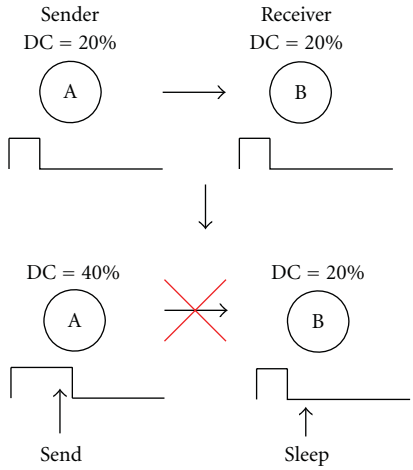


Figure 5: The packet throughput cannot be increased even if the node A increases its duty cycle because node B is in sleeping state.

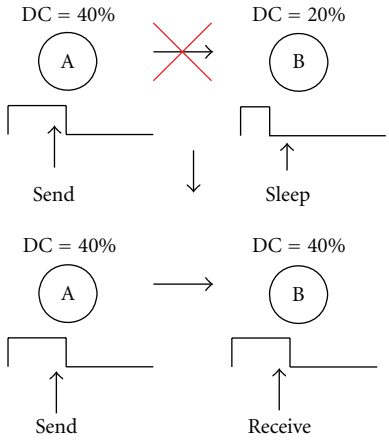


Figure 6: Node A and node B have the same duty cycle after duty cycle synchronization.

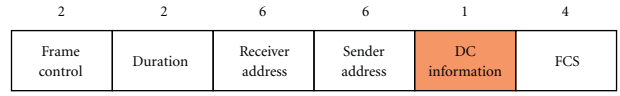


Figure 7: The data structure of the DCAdjust control packet.

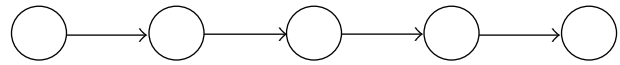


Figure 8: Linear topology.

Table 3: Common parameters of different MAC protocols.

Queue = 100
Sleep power = 0.02 mA
Transmission power = 10 mA
Receive power = 4 mA
Simulation time = 300 s
Bandwidth = 115.2 Kbps
Data packet size = 100 byte
SYNC period = 10 s
Frame time = 1 s
Initial duty cycle = 10%
DC_{max} = 40%
DC_{min} = 10%
Transmission range = 20 M
Node = 5~20

Table 4: Unique parameter of U-MAC and TA-MAC.

U-MAC		TA-MAC	
U_{high}	0.08	WR_{high}	8%
U_{low}	0.04	WR_{low}	4%
D_{max}	1.8	PL	2

are evaluated in the simulation: S-MAC, U-MAC, and TA-MAC. In the S-MAC control scheme, a frame length is set to one second and with a length of the active time 100 ms. The topologies used to evaluate our proposed scheme are shown in Figures 8 and 9. Figure 8 shows five nodes that transmit data in a linear topology. Figure 9 shows five nodes that transmit data in a cross topology.

In the simulation, the power consumption and data communication parameters are based on the value of TR1001 hybrid radio transceiver [13]. In our experiments, we set the transmission range to 20 meters. The bandwidth is set to 115.2 kbps. Each data packet size is 100 bytes and 4 bytes for control packets including RTS, CTS, and SYNC/LSYNC. The simulated parameters can be divided into two categories: (i) common parameters and (ii) unique parameters. The common parameters are used in all simulated MAC protocols. The unique parameters are used in specific MAC protocols for specifying special parameters to specific MAC protocols. Table 3 depicts the common parameters, and Table 4 depicts the unique parameter of different MAC protocols in the simulating environment.

Figures 10, 11, and 12 show the simulation results of the different MAC protocols in linear topology. Figure 10 shows the data throughput of the different MAC protocols. In Figure 10, TA-MAC and U-MAC increase sensors' duty cycle with n% under high traffic load when the sending/receiving queue is not full. When the sending queue usage exceeds the queue threshold ($Queue_{high}$), TA-MAC rapidly adjusts its duty cycle to the maximum duty cycle (DC_{max}) and sends a DCAdjust packet to notify the receiver. In TA-MAC, the sender can have more time to transmit data. In U-MAC, the duty cycle is in shortage due to U-MAC increases its duty cycle slowly with n% in such a condition. S-MAC adopts fixed duty cycle design to transmit data packets whether the traffic load is high or not. As a result, TA-MAC has better data throughput than U-MAC and S-MAC.

Figure 11 shows the transmission delay of the different MAC protocols in linear topology. The lack of working time in S-MAC and U-MAC increases the transmission latency between senders and receivers. TA-MAC can increase duty cycle significantly when the traffic load is high. Meanwhile,

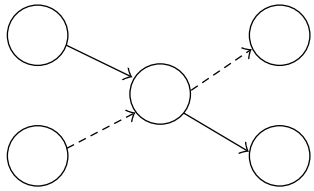


Figure 9: Cross topology.

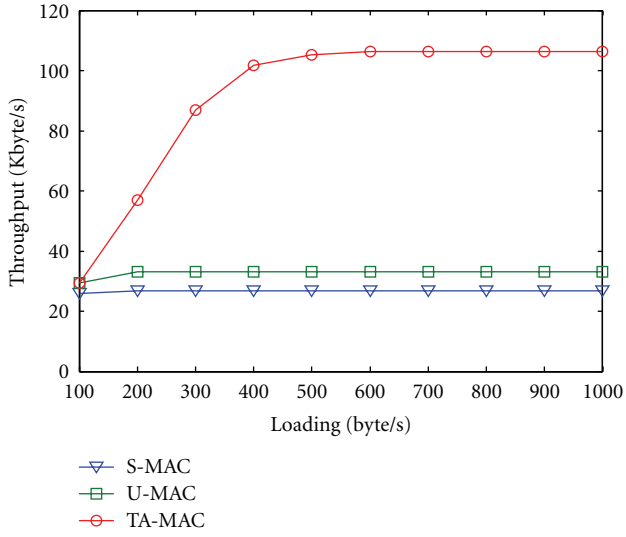


Figure 10: Data throughput for nodes in linear topology.

the sender and receiver can achieve duty cycle synchronization. Therefore, TA-MAC can have lower transmission latency than that in U-MAC and S-MAC.

Figure 12 shows the total energy consumptions of the different MAC protocols in linear topology. S-MAC uses a fixed duty cycle to transmit data. Therefore, the energy consumption of S-MAC does not increase as the traffic load increases. U-MAC increases sensor's duty cycle according to sensor's traffic load; however, the increasing is slow. Therefore, the energy consumption of U-MAC is increasing slowly until it reaches the maintenance level. TA-MAC adjusts its duty cycle to the maximum duty cycle (DC_{max}) and sends a DCAdjust packet to notify the receiver for improving data throughput, which consumes more battery energies. Therefore, TA-MAC takes more energy consumption than that in S-MAC and U-MAC for increasing data throughput and reducing transmission latency. In linear topology, the proposed TA-MAC can have good performance than S-MAC and U-MAC because TA-MAC can adjust the sensor's duty cycle efficiently. Meanwhile, the DCAdjust packet can make the sender and receiver have the same duty cycle, which can reduce the energy wasted in idle listening.

In the cross topology, the forwarding node suffers from high traffic loads from senders. The packet queue becomes full immediately when senders send data. Figure 13 shows the data throughput of the different MAC protocols in cross topology. In Figure 13, TA-MAC adjusts the forwarding node's duty cycle with n% according to its traffic load and sends a DCAdjust packet to the receiver for synchronizing

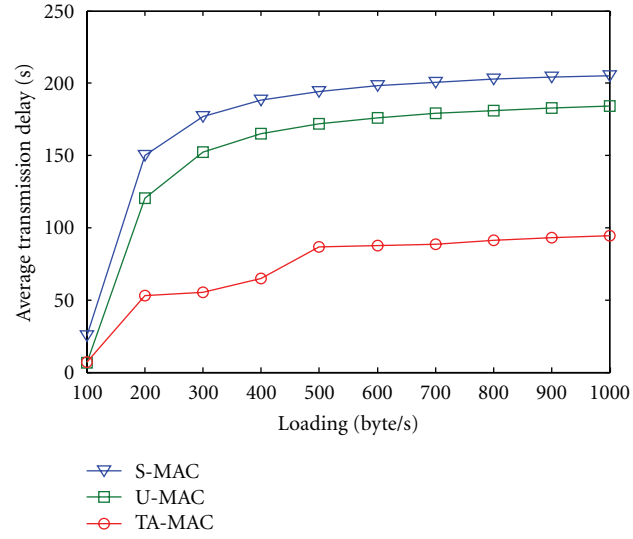


Figure 11: Transmission latency for nodes in linear topology.

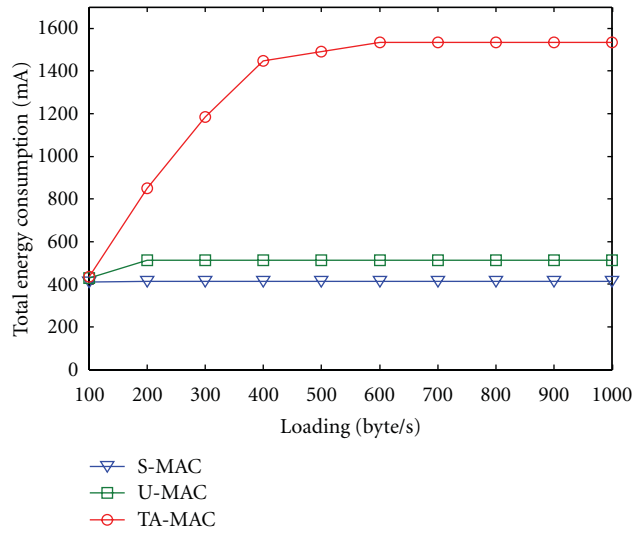


Figure 12: Energy consumption for nodes in linear topology.

duty cycle. When the receiver receives the DCAdjust packet, the receiver adjusts its duty cycle to the same value of the sender. After the duty cycle is synchronized, the sender can transmit data to the receiver without scheduling miss and then the data throughput is increasing rapidly. As the traffic load increases, the duty cycle of the forwarding node reaches the maximum duty cycle (DC_{max}) value. Then, the data throughput will be decreased due to the forwarding node cannot get more time to transmit data. S-MAC uses fixed duty cycle design to transmit data packets whether the traffic load is high or not. In S-MAC, the data throughput remained in a stable state. In U-MAC, the forwarding node increases its duty cycle with n% in such a condition. However, the duty cycles in sender and receiver are not synchronized. Therefore, the forwarding node cannot transmit data to the receiver due to scheduling miss. As a result, TA-MAC has better data throughput than U-MAC and S-MAC.

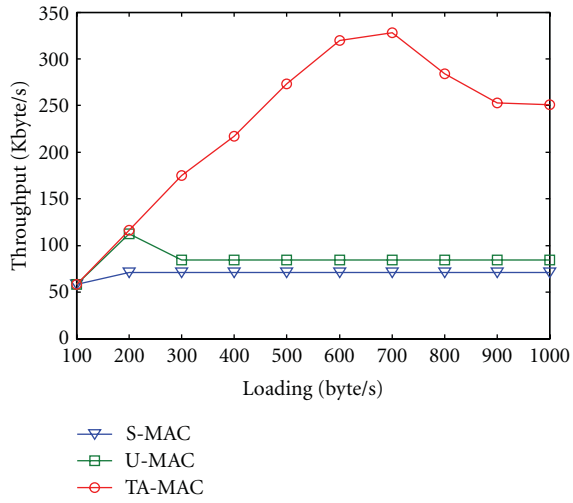


Figure 13: Data throughput for nodes in cross topology.

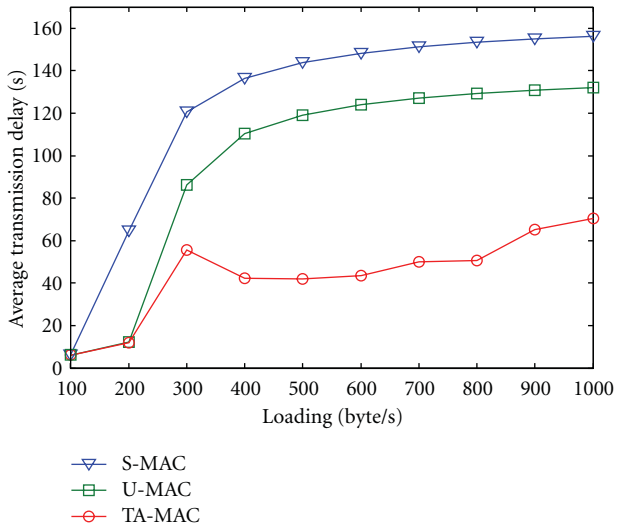


Figure 14: Transmission latency for nodes in cross topology.

Figure 14 shows the transmission delay of the different MAC protocols in cross topology. U-MAC adjusts the sensor's duty cycle according to its traffic load. U-MAC has longer duty cycle than that in S-MAC. Therefore, U-MAC has shorter transmission delay, compared with S-MAC. TA-MAC can increase duty cycle significantly when the traffic load is high. Meanwhile, the sender and receiver can achieve additional duty cycle synchronization. As a result, TA-MAC can have lower transmission latency than that in U-MAC and S-MAC.

Figure 15 shows the total energy consumptions of the different MAC protocols in cross topology. The energy consumption of S-MAC does not increase as the traffic load increases due to its fixed duty cycle design. U-MAC increases a sensor's duty cycle according to the sensor's traffic load. The forwarding node consumes more energy than other nodes. U-MAC has lower energy consumption than TA-MAC because the receivers in U-MAC are sleeping more due to out of duty cycle synchronization. On the contrast, T-MAC has more time to transmit queued data. Therefore, the energy consumption in TA-MAC is higher than that in S-MAC and U-MAC.

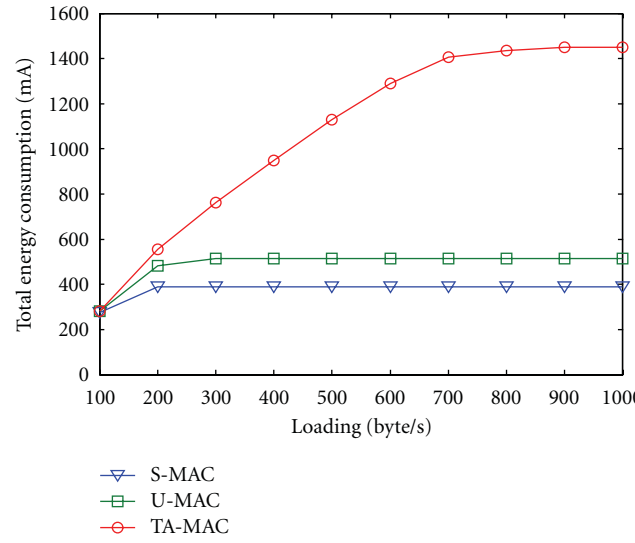


Figure 15: Energy consumption for nodes in cross topology.

5. Conclusions

In order to reduce the energy consumptions when nodes are in idle listening, duty-cycle-based MAC protocols are introduced to let node go into sleep mode periodically or aperiodically. When a sensor's listening time is extended, the sensor can have more time to receive data, which can enhance the transmission efficiency and reduce the latency when sensors have the heavy load. When a sensor's duty cycle is reduced, the sensor has long sleep time to reduce energy consumption in idle listening. In this paper, a dynamic traffic-aware MAC protocol for energy conserving in wireless sensor networks is proposed. The proposed MAC protocol can provide better data transmission rate when sensors are with high traffic loading. Meanwhile, the proposed MAC protocol can save energy when sensors are with low traffic loading. Simulation results show that the proposed protocol has better data throughput than other duty-cycle-based MAC protocols, for example, S-MAC and U-MAC.

Disclosure

C.-C. Chen was an associate professor in Southern Taiwan University while this work was initiated, and he is with Institute of Manufacturing Information and Systems, National Cheng Kung University since 2011.

References

[1] 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 (HICSS '00)*, pp. 3005–3014, January 2000.

[2] J. Li and G. Y. Lazarou, "A bit-map-assisted energy-efficient MAC scheme for wireless sensor networks," in *Proceedings of the 3rd International Symposium on Information Processing in Sensor Networks (IPSN '04)*, pp. 55–60, April 2004.

[3] Y. S. Chen and Y. W. Lin, "C-MAC: an energy-efficient MAC scheme using chinese-remainder-theorem for wireless sensor networks," *Journal of Information Science and Engineering*, vol. 23, no. 4, pp. 1057–1071, 2007.

[4] C. Shanti and A. Sahoo, "DGRAM: a delay guaranteed routing and MAC protocol for wireless sensor networks," in *Proceedings of the 9th IEEE International Symposium on Wireless, Mobile and Multimedia Networks (WoWMoM '08)*, June 2008.

[5] H. Gong, M. Liu, L. Yu, and X. Wang, "An event driven TDMA protocol for wireless sensor networks," in *Proceedings of the WRI International Conference on Communications and Mobile Computing (CMC '09)*, vol. 2, pp. 132–136, January 2009.

[6] L. Shi and A. O. Fapojuwo, "TDMA scheduling with optimized energy efficiency and minimum delay in clustered wireless sensor networks," *IEEE Transactions on Mobile Computing*, vol. 9, no. 7, Article ID 5432185, pp. 927–940, 2010.

[7] D. Wu, G. -Y. Wang, and X. -L. Li, "Distributed TDMA scheduling protocol based on conflict-free for wireless sensor networks," in *Proceedings of the International Conference on Intelligent Computing and Integrated Systems (ICISS '10)*, pp. 876–879, October 2010.

[8] W. Zhao and X. Tang, "Scheduling data collection with dynamic traffic patterns in wireless sensor networks," in *Proceedings of the International Conference on Computer Communications (IEEE INFOCOM '11)*, pp. 286–290, April 2011.

[9] W. Ye, J. Heidemann, and D. Estrin, "An energy-efficient MAC protocol for wireless sensor networks," in *Proceedings of the IEEE International Conference on Computer Communications (IEEE INFOCOM '02*, vol. 3, pp. 1567–1576, June 2002.

[10] T. van Dam and K. Langendoen, "An adaptive energy-efficient MAC protocol for wireless sensor networks," in *Proceedings of the 1st International Conference on Embedded Networked Sensor Systems (SenSys '03)*, pp. 171–180, November 2003.

[11] S. H. Yang, H. W. Tseng, E. H. K. Wu, and G. H. Chen, "Utilization based duty cycle tuning MAC protocol for wireless sensor networks," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '05)*, vol. 6, pp. 3258–3262, December 2005.

[12] J. Polastre, J. Hill, and D. Culler, "Versatile low power media access for wireless sensor networks," in *Proceedings of the Second International Conference on Embedded Networked Sensor Systems (SenSys '04)*, pp. 95–107, November 2004.

[13] A. El-Hoiydi and J. D. Decotignie, "Low power downlink MAC protocols for infrastructure wireless sensor networks," *Mobile Networks and Applications*, vol. 10, no. 5, pp. 675–690, 2005.

[14] M. Buettner, G. Yee, E. Anderson, and R. Han, "X-MAC: a short preamble mac protocol for duty-cycled wireless networks," Tech. Rep. CU-CS-1008-06, University of Colorado at Boulder, 2006.

[15] A. El-Hoiydi, "Aloha with preamble sampling for sporadic traffic in Ad Hoc wireless sensor networks," in *Proceedings of the International Conference on Communications (ICC '02)*, pp. 3418–3423, April 2002.

[16] P. Lin, C. Qiao, and X. Wang, "Medium access control with a dynamic duty cycle for sensor networks," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '04)*, vol. 3, pp. 1534–1539, March 2004.

[17] G. Xing, C. Lu, Y. Zhang, Q. Huang, and R. Pless, "Minimum power configuration in wireless sensor networks," in *Proceedings of the 6th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MOBIHOC '05)*, pp. 390–401, May 2005.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 514836, 12 pages
doi:10.1155/2012/514836

Research Article

Design of an Effective WSN-Based Interactive u-Learning Model

Hye-jin Kim,[1] Ronnie D. Caytiles,[2] and Tai-hoon Kim[3]

[1] *Continuing Education Center, Jeonju University, Jeonju 560-759, Republic of Korea*
[2] *Multimedia Engineerig Department, Hannam University, Daejeon 306-791, Republic of Korea*
[3] *GVSA and University of Tasmania, Newnham, TAS 7248, Australia*

Correspondence should be addressed to Tai-hoon Kim, taihoonn@paran.com

Received 17 September 2011; Accepted 17 October 2011

Academic Editor: Sabah Mohammed

Wireless sensor networks include a wide range of potential applications to improve the quality of teaching and learning in a ubiquitous environment. WSNs become an evolving technology that acts as the ultimate interface between the learners and the context, enhancing the interactivity and improving the acquisition or collection of learner's contextual information in ubiquitous learning. This paper presents a model of an effective and interactive ubiquitous learning environment system based on the concepts of ubiquitous computing technology that enables learning to take place anywhere at any time. The u-learning model is a web-based e-learning system utilizing various state-of-the-art features of WSN that could enable learners to acquire knowledge and skills through interaction between them and the ubiquitous learning environment. It is based on the theory of connectivism which asserts that knowledge and the learning of knowledge are distributive and are not located in any given place but rather consist of the network of connections formed from experiences and interactions with a knowing community. The communication between devices and the embedded computers in the environment allows learners to learn in an environment of their interest while they are moving, hence, attaching them to their learning environment.

1. Introduction

Learning is everywhere. It is not only found inside the four corners of the classroom. It can be in homes, workplaces, playgrounds, libraries, and even in our daily interactions with others. People can learn anywhere and in any time convenient for them without any hassles to their routine works and activities. There are new things that do not need to be learned in formal classrooms, these can be learned at user-friendly interfaces. And people are eager to become educated, professional per se, even without attending formal school, and a new innovation in educational technology would be paramount to their goals, aims, and needs.

Ubiquitous learning is a new educational learner-centered paradigm characterized by providing intuitive ways for identifying the right collaborators, the right contents, and the right learning services in the right place and the right time based on the student's surroundings. That is, who are the learning collaborators that could provide the student's needs, what are the learning resources and services available, and when and where should the learning take place?

The ubiquitous learning environment utilizes a large number of cooperative small nodes with computing and communication facilities such as handheld terminals, smart mobile phones, sensor network nodes, contactless smart cards, RFID (radio frequency identification), and mobile IP.

In this paper, a ubiquitous learning environment is created based on the concepts of ubiquitous computing technology that enables learning to take place anywhere at any time. Various e-learning environments, architectural designs, and implementations are examined and considered to include the elements of m-learning and wireless sensor networks (WSNs), extending them to create a new u-learning environment.

The ubiquitous learning environment guided by principles of learning paradigms advances the computing capabilities allowing educators to become catalysts of learning, transforming students into a more active and collaborative participants in knowledge making. By modifying the traditional

role of education to embrace principles of connectivism as applied to the described u-learning environment, education can respond to and address the onset digital revolution for future development of ubiquitous learning.

This new ubiquitous learning environment is described as an environment that supports student's learning using digital media in geographically distributed environments. This learning environment is the basis for the design of an effective WSN Based Interactive u-Learning model where learning takes place between students within campus/home and teachers as facilitators in u-space.

The u-learning model is a web-based e-learning system utilizing various state-of-the-art features of WSN that could enable learners to acquire knowledge and skills through interaction between them and the defined ubiquitous learning environment. Students are allowed to be in an environment of their interest wherein the communication between devices and the embedded computers in the environment allows learners to learn while they are moving, hence, attaching them to their learning environment.

2. Relevant Researches

The beneficial effects of learning enhanced by technology have been discussed and demonstrated by researchers in the past decades. The use of technology can make learning easier and more effective that could provide a lifeline for isolated and helpless learners.

The benefits and features of applying the context-aware ubiquitous learning which includes the provision of a more adaptive and active learning supports, the integration of real-world and digital learning scenarios, and the accomplishment of real-world practice environments with portfolio-recording functions have been recently indicated and focused on by researchers [1–4].

The following related researches have been dealing with the technical aspect of ubiquitous learning focused on the use of context-aware approach and wireless sensor networks.

To develop context-aware and seamlessly integrated Internet environments, a variety of new techniques and products concerning ubiquitous computing have been developed in recent years, such as sensors and actuators, RFID (radio frequency identification) tags and readers, wireless communication, mobile phones, PDAs, and wearable computers [5, 6].

In [7], anyone can make use of computers that are embedded everywhere in a public environment at any time in a ubiquitous computing environment. A user equipped with a mobile device can connect to any computer and access the network using wireless communication technology [8]. Moreover, not only can a user access the network actively, but computers around the user can recognize their behavior and offer various services according to their situation, the mobile terminal's facilities, and the network bandwidth. User assistance via ubiquitous computing technologies is realized by providing users with proper decisions or decision alternatives. That is, a ubiquitous computing technology-equipped system supplies users with timely information and relevant services by automatically sensing their various context data and effectively generating the proper results [9].

The attention of researchers has been attracted by the advent of u-computing techniques from the fields of both education and computer science.

Developing context-aware toolkits which can provide functionalities that enable adaptive service based on personalized contexts were also given attention in the recent years [10]. Also, several context models have been proposed to record and analyze user behaviors in the real world such that high-quality service can be provided [11]. The use of RFID tags and floor-mounted weight sensors to detect the spatiotemporal relationship between a human user and various objects has been demonstrated. A model that represents the user's state, in which people's activities are described in terms of time sequence aspect in addition to location aspect, was proposed.

The next set of related researches have been dealing with the pedagogical aspects of ubiquitous learning focused on the use of context-aware approaches and theories of connectivity.

Based on surveys and researches, most students prefer "authentic activities" in which they can work with problems from the real world [12]. The importance of providing necessary "scaffolding" for novices to operate within the complex realistic context has been emphasized. Moreover, the need for the provision of supports that enable teachers to track progress, assess information, interact knowledgeably and collaboratively with individual students or cooperating groups of students, and prepare situated learning activities to assist the students in improving their ability in utilizing skills or knowledge has also been recognized [13].

Context-aware u-learning is an approach that places the students in a series of designed lessons that combine both real and virtual learning environments. The connection between learner-centered and real world-situated learning has been clearly revealed in the relevant research.

Context-aware u-learning has provided an adequate environment with cognitive apprenticeships. The features of cognitive apprenticeships include situated learning, coaching, scaffolding, reflection, and exploration. That is, a learning system of applying cognitive apprenticeships should be situated in real-life context. Also, it should provide systematic teaching and guidance for the learners, and opportunities of practicing learning tasks as well as reviewing learning processes.

Context-aware u-learning is able to provide personalized and active support to assist students to learn in the real world, which is very important from both the learning attitude and the learning effectiveness aspect. Also, it can deliver the provision of more opportunities for practicing and the saving of manpower in assisting and monitoring the learners.

3. Background/Rationale

3.1. What Is u-Learning? u-learning involves creating learning activities, tasks, projects, and resources that encourage students to discover learning themselves without consciously

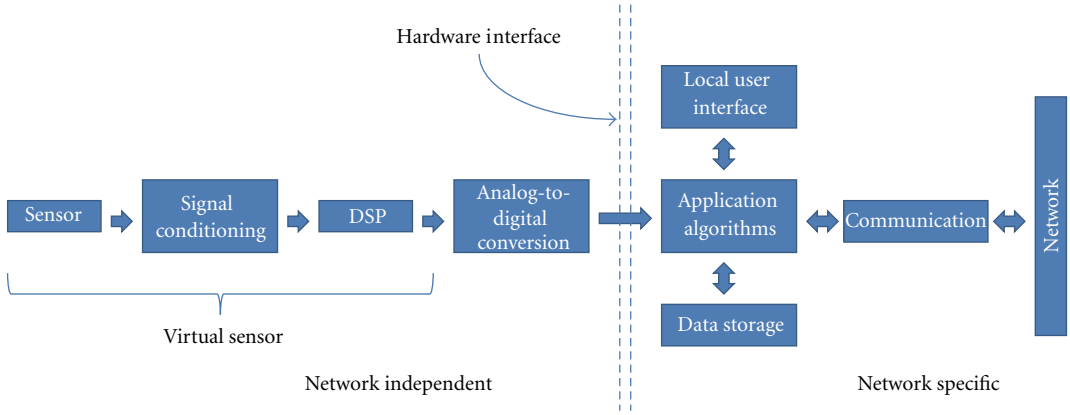


Figure 1: Sensor model.

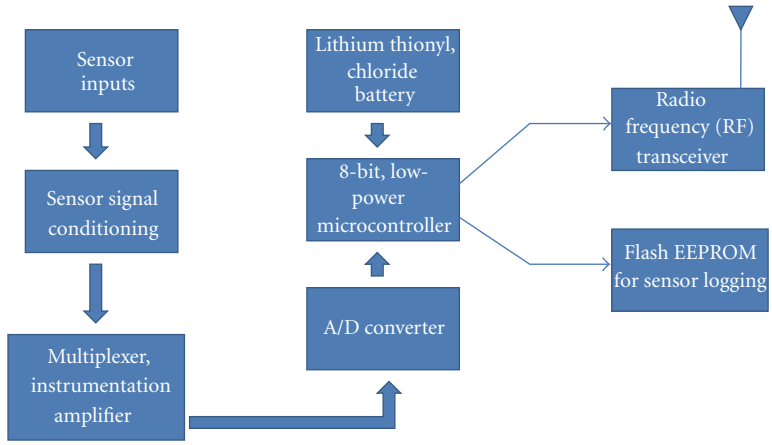


Figure 2: Wireless sensor node functional block diagram.

realizing they are learning, so that they learn automatically and independently.

"Ubiquitous learning is ideas without barriers, inspiration without limits, innovation without boundaries." "Ubiquitous learning requires no less than a fundamental reinvention of the science of education for our modern society" [14].

Tony Stockwell author of Accelerated Learning in Theory and Practice and The Learning Revolution says, "To learn anything fast and effectively, you have to see it, hear it and feel it." This means the learners need to have mental links (pictures, sound, videos, actions, and so forth) to help them visualize learning fast and effectively [15].

u-learning system is an environment supporting student learning using digital media in a geographically distributed environment. This environment proposed the implementation of student learning between students within campus/at home and teacher of content producer in u-space, which is not limited to traditional learning system [16, 17].

Ubiquitous learning is a learning style in which the learner can smoothly commence the learning process anytime, anywhere [18].

3.2. What Is Wireless Sensor Networks (WSNs)? Sensor networks are the key to gathering the information needed by smart and ubiquitous environments, whether in buildings, utilities, industrial, home, shipboard, transportation systems, automation, education, or elsewhere. Figure 1 depicts the model of a smart wireless sensor.

A wireless sensor network (WSN) consists of spatially distributed autonomous sensors to monitor physical or environmental conditions, such as temperature, sound, vibration, pressure, motion, or pollutants and to cooperatively pass their data through the network to a main location. The more modern networks are bidirectional, enabling also to control the activity of the sensors. WSN development was initially motivated by military applications such as battlefield surveillance; today such networks are used in many industrial and consumer applications, such as industrial process monitoring and control and machine health monitoring.

Wireless sensors are networked and scalable, consume very little power, are smart and software programmable, capable of fast data acquisition, reliable and accurate over the long term, cost little to purchase and install, and require no real maintenance [19]. They generally consist of a base station (or "gateway") that can communicate with a number of wireless sensors via a radio link. Data is collected at the wireless sensor node, compressed, and transmitted to the gateway directly or, if required, uses other wireless sensor nodes to forward data to the gateway. The transmitted data is then presented to the system by the gateway connection. Figure 2 depicts the functional block diagram of a wireless sensor node.

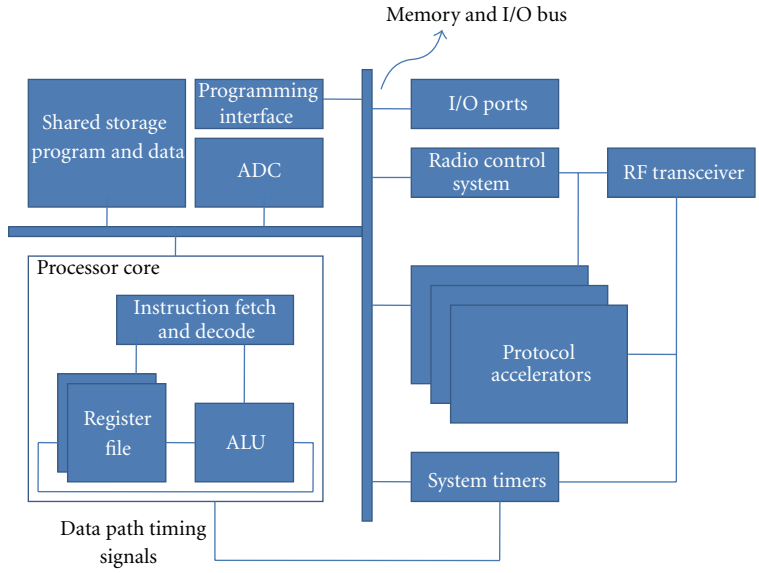


Figure 3: Wireless sensor architecture.

WSNs can be deployed anywhere with the IEEE802.15.4 as its communication protocol [20], and utilize star and mesh topologies enhancing the dynamic of such type of networks. Each device is called a mote or sensor node and each network has one powerful node called Sink Node. The Sink Node is the interface to the exterior of the network. Each mote is capable of sensing parameters such as light, temperature, humidity, pressure, acceleration, and position.

Figure 3 describes the general architecture for an embedded wireless sensors device.

3.3. Mobile IP. Mobile IP is a standard that allows users with mobile devices whose IP addresses are associated with one network to stay connected when moving to a network with a different IP address. When a user leaves the network with which his device is associated (home network) and enters the domain of a foreign network, the foreign network uses the mobile IP protocol to inform the home network of a care-of address to which all packets for the user's device should be sent [21].

Figure 4 describes the general operation of a mobile IP protocol. The mobile IP protocol allows location-independent routing of IP datagrams on the Internet. Each mobile node is identified by its home address disregarding its current location. While away from its home network, a mobile node is associated with a care-of address which identifies its current location, and its home address is associated with the local endpoint of a tunnel to its home agent. Mobile IP specifies how a mobile node registers with its home agent and how the home agent routes datagrams to the mobile node through the tunnel [22].

Mobile IP is most often found in wireless WAN environments where users need to carry their mobile devices across multiple LANs with different IP addresses.

Mobile IPv6 is a version of mobile IP—a network layer IP standard used by electronic devices to exchange data across a packet switched internetwork. Mobile IPv6 allows an IPv6 node to be mobile—to arbitrarily change its location on an IPv6 network—and still maintain existing connections [22]. IPv6 (Internet Protocol version 6) is the successor to the well-known IPv4 protocol, commonly known as IP.

3.4. The Theory of Connectivism. The learning theory flows through communication theory (communication consists of information that flows through a channel) that is, information (audio music and video images, for example) is transmitted from a sender to a receiver. The effectiveness of communication is improved through interaction. The viewing of communication should not be done only as a one-time event, in which information is sent by teachers and received by the learners; the transfer of information is enabled through a series of communications, such that the receiver sends messages back to the sender, or to third parties.

3.4.1. The Network Analogy. The process of learning as a whole operates more like a social network than a digital computer. The analogy with the network of computers is a good example in which the actual social network itself—a set of distributed and interlinked entities, usually people, as represented by websites or pages—constitutes a type of distributed representation. Specific mental operations, therefore, are like thinking of functions applied to this social network. Based on these mental operations, the following implications are derived: knowledge is subsymbolic (mere possession of the words does not mean that there is knowledge; the possession of knowledge does not necessarily result in the possession of the words); knowledge is distributed across a network of people; knowledge is interconnected; knowledge is personal (self-organized); knowledge is an emergent phenomenon.

Networks consist of three major elements: entities—the things that are connected that send and receive signals;

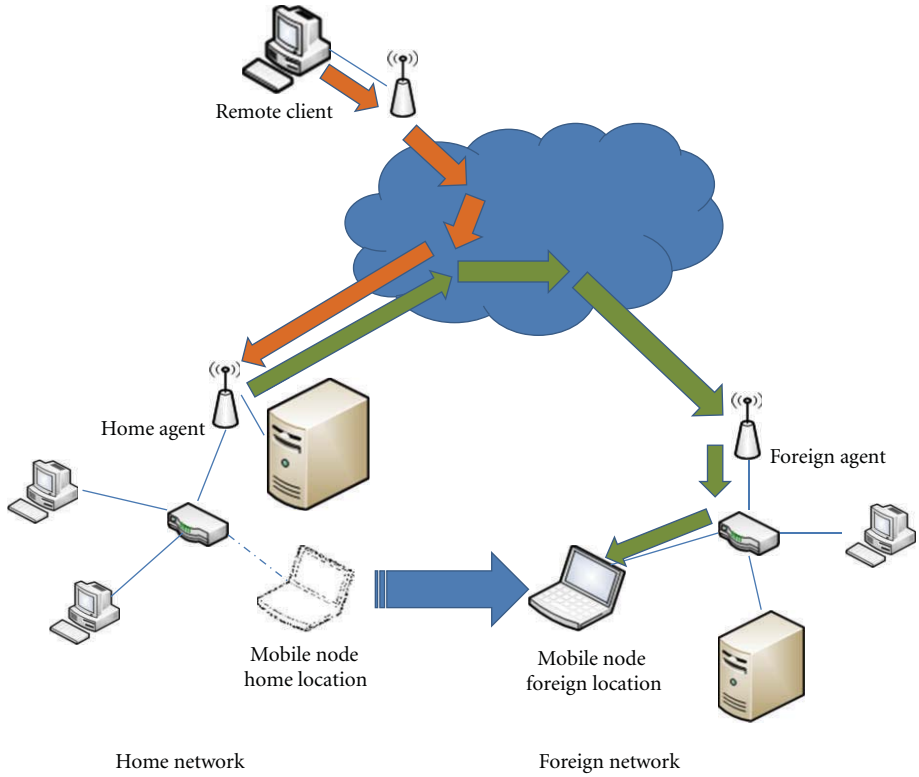


Figure 4: Mobile IP protocol general operation.

connections—the link or channel between entities; signals—the message sent between entities.

A network is defined as connections between entities. Computer networks, power grids, and social networks all function on the simple principle that people, groups, systems, nodes, and entities can be connected to create an integrated whole. Alterations within the network have ripple effects on the whole.

Networks are classified according to the following set of properties: density—the number of entities the other entity is connected to; speed—the measure of how long a message moves from entity to entity; flow—the measure of how much information an entity processes, which includes messages sent and received in addition to transfers of messages for other entities; plasticity—the measure of how frequently connections are created or abandoned; the degree of connectedness—a function of density, speed, flow, and plasticity.

The objective of learning the theory of networks is to describe the manner in which resources and services are organized in order to offer learning opportunities in a network environment. Learning networks are not therefore a pedagogical principle, but rather, a description of an environment intended to support a particular pedagogy.

Social network analysis is an additional element in understanding learning models in a digital era. Within social networks, hubs are well-connected people who are able to foster and maintain knowledge flow. Their interdependence results in effective knowledge flow, enabling the personal understanding of the state of activities organizationally.

3.4.2. What Is Connectivism? Connectivism is the theory that knowledge is distributed across a network of connections, and therefore that learning consists of the ability to construct and traverse those networks. It shares with some other theories a core proposition, that knowledge is not acquired, as though it were a thing. Knowledge is, in this theory, literally the set of connections formed by actions and experience [23].

From the article of George Siemens [24], connectivism is the integration of principles explored by chaos, network, and complexity and self-organization theories. Learning is a process that occurs within nebulous environments of shifting core elements—not entirely under the control of the individual. Learning (defined as actionable knowledge) can reside outside of ourselves (within an organization or a database), is focused on connecting specialized information sets, and the connections that enable us to learn more are more important than our current state of knowing. Connectivism is driven by the understanding that decisions are based on rapidly altering foundations. New information is continually being acquired. The ability to draw distinctions between important and unimportant information is vital. The ability to recognize when new information alters the landscape based on decisions made yesterday is also critical.

The following are the principles of connectivism as discussed by Siemens.

(i) Learning and knowledge rests in diversity of opinions.

(ii) Learning is a process of connecting specialized nodes or information sources.

(iii) Learning may reside in nonhuman appliances.

(iv) Capacity to know more is more critical than what is currently known.

(v) Nurturing and maintaining connections are needed to facilitate continual learning.

(vi) Ability to see connections between fields, ideas, and concepts is a core skill.

(vii) Currency (accurate, up-to-date knowledge) is the intent of all connectivist learning activities.

(viii) Decision-making is itself a learning process. Choosing what to learn and the meaning of incoming information is seen through the lens of a shifting reality. While there is a right answer now, it may be wrong tomorrow due to alterations in the information climate affecting the decision.

Connectivism also addresses the challenges that many corporations face in knowledge management activities. Knowledge that resides in a database needs to be connected with the right people in the right context in order to be classified as learning.

Information flow within an organization is an important element in organizational effectiveness. In a knowledge economy, the flow of information is the equivalent of the oil pipe in an industrial economy. Creating, preserving, and utilizing information flow should be a key organizational activity. Knowledge flow can be likened to a river that meanders through the ecology of an organization. In certain areas the river pools, and in other areas it ebbs. The health of the learning ecology of the organization depends on effective nurturing of information flow.

4. Interactive WSN-Based u-Learning Model

4.1. What Is an Interactive u-Learning Model? u-learning is a learning paradigm which takes place in a ubiquitous computing environment that enables learning the right thing at the right place and time in the right way. It is an expansion of previous learning paradigms as we move from conventional learning to electronic learning to mobile learning and now to ubiquitous learning (thus, the meaning of "e" is not just limited to "electronic" but expands to "everywhere," "extending", "enhancing," and "enabling."

u-Learning enables us to change our current learning processes to be more efficient and more effective. If done right, u-learning becomes a critical force to improve the performance of our workforce and our organization as a whole [25].

Table 1: WSN Measurable items.

Quantity	Description
Sound	Sound intensity and velocity recognition
Temperature	Heat measurement, source recognition
Light	light intensity measurement, source recognition
Air	Air quality identification, different types of gases identification
Water	Water components, minerals, levels, source identification
Humidity	Humidity measurement

4.2. Wireless Sensor Networks for u-Learning. Educational activities when integrated with wireless sensor networks can help learners improve learning, by becoming more active and making it an easier and more enthusiastic and interactive activity. Learners can interact with the real context obtained from WSNs which could help enhance understanding and learning in specific areas. Table 1 depicts some of the quantities that can be measured and evaluated using WSNs.

WSNs can show learners the different measurements, values, and other characteristics of some parameters of the environment. Learners can be able to observe, collect, and experiment with these parameters with enhanced excitement, and interactivity. For instance, the study of properties like sound, temperature, light, or humidity can be easily supported using WSNs.

Figure 5 shows the infrastructure for the deployment of various Wireless Sensor Networks described in Table 1. The data collection utilizes Zigbee or 802.15.4/6LoWPAN lows-powered mesh topology which is then transmitted to a secured IPs-based infrastructure. WSNs make learners interact with everything in their context, anytime anywhere (context-aware learning triggered by the learner).

4.3. Connectivism for u-Learning Model. The starting point of connectivism is the individual learner. Personal knowledge is comprised of a network, which feeds into organizations and institutions, which in turn feed back into the network, and then continue to provide learning to individual. This cycle of knowledge development (personal to network to organization) allows learners to remain current in their field through the connections they have formed.

The amplification of learning, knowledge, and understanding through the extension of a personal network is considered as the epitome of connectivism.

The following are the implications of connectivism to the u-learning model.

(i) Management and leadership. The management and marshalling of resources to achieve desired outcomes are a significant challenge.

(ii) Innovation: diverse teams of varying viewpoints become the critical structure for completely exploring ideas.

Figure 5: WSN deployment with IP support for interactive u-learning.

(iii) An organization's ability to foster, nurture, and synthesize the impacts of varying views of information is critical to knowledge economy survival, thus, the speed of "idea to implementation" is improved in a systems view of learning.

(iv) The trend of media, news, and information are being challenged by the open, real-time, two-way information flow of blogging.

(v) Personal knowledge management in relation to organizational knowledge management.

(vi) The design of the learning environments.

The theory of connectivism describes a model of learning that acknowledges the tectonic shifts in society where learning is no longer an internal, individualistic activity. How people work and function is altered when new tools are utilized. The field of education has been slow to recognize both the impact of new learning tools and the environmental changes in what it means to learn. Connectivism provides insight into learning skills and tasks needed for learners to flourish in a digital era.

4.4. Proposed Interactive u-Learning Model. u-learning as an interactive social learning model is designed in consideration of some factors that mainly influence the learning process of a learner—member of the net generation. Management systems and structures, competencies, culture, and technology have a direct impact on the learning process of the learners. The media influence in learning plays a big role in the advancements of the learning processes. Media evolved from print to nonprint that could direct attention, arouse motivation, increase student's concentration, and help them to be actively involved in the learning process [26] (Figure 6).

The advancement of computing technologies together with the enhancement of wireless communication technologies nowadays help out to support the expansion of ubiquitous computing, hence, lead to the advancements of the learning processes. A variety of computing and communication technologies have been developed, such as

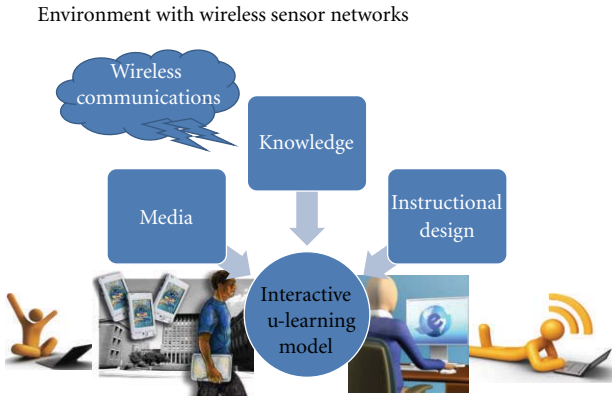


Figure 6: WSN-based interactive u-learning model.

sensors and actuators, RFID (radio frequency identification) tags and cards, wireless communication equipment, mobile phones, PDAs (personal digital assistants), and wearable computers.

This model aims for accommodating learners in their learning style by providing adequate information anytime and anywhere as they wish for it. To promote a more effective application of u-learning, a model for the design of an online learning environment is presented. This model guided by principles of learning paradigms advances the computing capabilities allowing educators to become catalysts of learning, transforming students into more active and collaborative participants in knowledge making.

Knowledge repository refers to a wide range of knowledge domains across various subject areas including math, physics, biology, social science, and language, it is being interacted upon by goal analysis, task analysis, learner characteristics, and other semantic rules to provide a design for knowledge representation for learners.

Media repository represents the different media to support the knowledge representation in u-learning. For example, learners can use multimedia tools to create a learning object that contains the new knowledge created by the learner and post it to the web to share with others. Media refer to the use of one medium (e.g., video) or a combination of several media (e.g., video, audio, textbook, etc.) in instruction.

Instructional design repository consists of all the design components in instructional design. This is the basis for the design and modality of the learning modules in the u-learning model.

The u-learning model described presents information in an interactive and informative way. This may include courses, student's information, and teacher's information. The development of the learning content is based on the instructional designs applied to the information stored on the knowledge repository.

Based on the inputs from the instructional design component and domain knowledge analyses in the knowledge repository, the knowledge representation is formed. Various learning modules are formulated to represent the knowledge identified, and at the same time the rules of reusability and sharability are applied to the design and development of instructional modules. Unlike the traditional design, knowledge representation enables learners to experience the construction and creation of knowledge through multiple venues, thus promoting a learning process that focuses on the understanding of "how do people know it" rather than "what do people know."

An online instructional program is created to facilitate dynamic knowledge acquisition and creation as well as to promote learners' self-initiation and collaboration in learning. In the design of the instructional program, the current state and needs of the learner are determined, the end goal of each instruction is defined, and a learning intervention to assist the acquisition of new skills, knowledge, or expertise is developed.

4.5. Proposed Interactive u-Learning Characteristics. The interactive u-learning model was designed and developed considering the two major areas.

(i) Resources—students needed one place to find both internal (forms, FAQs, course offerings, etc.) and external information (RSS feeds, bookmarks, etc.).

(ii) Community—students needed a place to collaborate, communicate, and share information such as ideas, photos, videos, and news.

Characteristics of an interactive u-learning model [12].

(i) Permanency: the information remains and learners can never lose their work unless the learners purposely remove it.

(ii) Accessibility: the information is always available whenever the learners need to use it. Learners have access to their documents, data, or videos from anywhere.

(iii) Immediacy: the information can be retrieved immediately by the learners wherever they are.

(iv) Interactivity: the learners can interact with peers, teachers, and experts efficiently and effectively through different media in the form of synchronous or asynchronous communication.

(v) Adaptability: learners can get the right information at the right place with the right way.

(vi) Context awareness: the environment can adapt to the learners, real situation to provide adequate information for the learners.

(vii) Theory based: contents of the learning modules are based on contemporary approaches to teaching and learning.

(viii) Innovative and relevant: contents of the learning modules are designed based on the learning objectives.

(ix) Emergent: allowing (where appropriate) the interactions between course participants and enable them to actively explore the relevance and application of the course content.

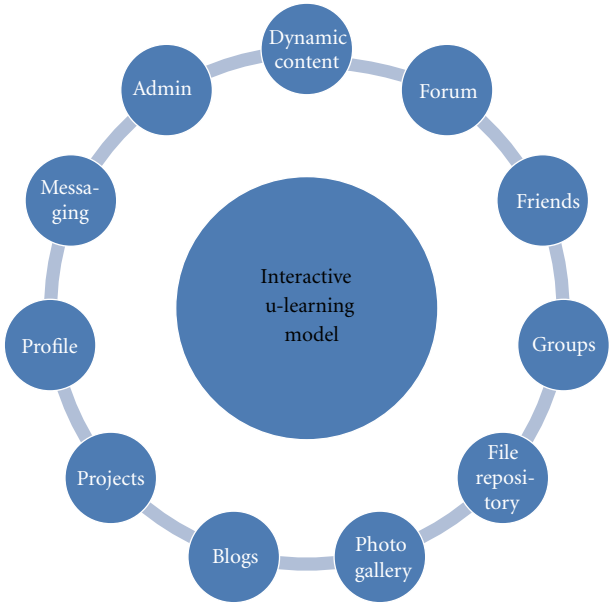


Figure 7: Interactive u-learning model functionalities.

(x) Personalized: participants are able to apply their own context and situation to the learning outcomes. The learning could be embedded in the learner's daily life.

Learners are offered the opportunity to increase the implements using the latest multimedia technology, equipment, and testing.

4.6. Proposed Interactive u-Learning Model Functionalities. Some of the major features and functionality of the interactive u-learning model include the following (Figure 7).

(i) Dynamic content—user-generated content can appear throughout the website in addition to automatically generated content such as RSS feeds of news, journals, and blogs.

(ii) Forums—registered users can create discussion threads about courses, research, administration, or just about anything.

(iii) Projects—registered users can create and manage all aspects of a project and invite others to join the project through an online project management tool. This function could be found on the user's profile.

(iv) Profile—registered users are automatically given a profile that can be edited and updated with personal and profession information.

(v) Messaging—registered users are allowed to send and receive private messages within the community.

(vi) Administration—allows admin users to maintain courses or instructional modules provided for students. This function is provided for teachers to allow them some administrative functions of the e-learning system.

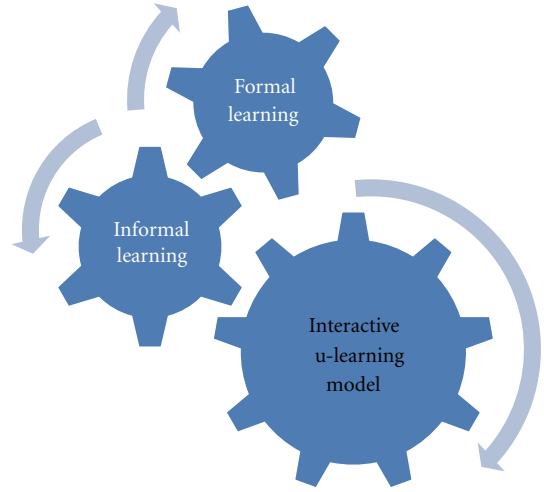


Figure 8: u-learning implementation model.

(vii) Friends—registered users can connect and create friendship networks by accepting and requesting friends within the u-learning community.

(viii) Groups—allows users with similar interests, or perhaps users as part of a research groups or a course/class, to have a more structured setting to share content and discuss ideas. Users can create and moderate as many groups as they like. Group activity could be kept private to the group or the "make public" option can be used to disseminate work to the wider public. Each group produces granular RSS feeds, so it is easy to follow group developments. Each group has its own URL and profile, and each group comes with a file repository, forum, pages, and a message board.

(ix) Blogs—users are allowed to publish entries of commentary, descriptions of events, or other material such as graphics or video.

(x) Photo gallery—allows users to manage and organize their photo collections by adding titles, rating, captions, and custom metadata tags to photos. Users can also see pictures that their friends have uploaded, or see pictures attached to a group. Clicking on an individual file shows a larger version of the photo.

(xi) File repository—allows users to upload any kind of file. The uploaded files can be filtered by tag and users can restrict access so that files are only visible to the people they want. Each file may also have comments attached to it.

4.7. u-Learning Implementation. u-Learning contents are developed based on two major areas as shown in Figure 8. One is the resources for the process of an academic study (formal learning). These include the academic calendars, the program descriptions and requirements, various forms for portfolio reviews, advance to candidacy, and graduation. The u-learning model provides one place to access all these

Figure 9: u-learning community implementation tools.

materials. WSNs were utilized to enhance mobility and interactivity and to create a new learning environment.

The other area is for the social needs of learners (informal learning), which can help to develop a learning community. These include access to forums, sites with RSS feeds, blogs, alerts, journals, and research threads, this is the place where learners collaborate, communicate, and share information such as ideas, photos, videos, and news.

The development of the u-learning web-based e-learning system is focused on open-source technology solutions with flexibility and scalability.

Some e-learning implementations are analyzed and the important features and elements needed to accommodate the learning process of the net generation learners are identified. Then, the different technology solutions are reviewed as suggested by the current trends in technology developments while looking at new open-source options available to determine the best solution for the development needs.

These technology options were chosen based on compatibility, having the required features, time and skill required for implementation, flexibility, ease of use, and future scalability (Figure 9).

It is suggested to use Joomla as the content management system (CMS) solution which has a strong administrative focus that allows nondevelopers to quickly install and manage the system. It also offered a large number of modules that can be easily installed to meet the needs of the community [27].

The integration to this content management solution of Moodle, an open source learning management system (LMS), allows even nontechnical teachers to set up and maintain where students can log in, access course information, interact, share, and teach others. Moodle's main focus and purpose are for managing learning activities and users, but it also has a built-in functionality for blogging, wikis, and many other applications similar to Joomla.

Moodle also known as course management system has several features typical of an e-learning platform, plus some original innovations (like its filtering system). Moodle can be used in many types of environments such as in education, training and development, and business settings [28].

Developers can extend Moodle's modular construction by creating plugins for specific new functionality. Moodle's infrastructure supports many types of plug-ins:

(i) activities (including word and math games),

(ii) resource types,

(iii) question types (multiple choice, true and false, fill in the blank, etc.),

(iv) data field types (for the database activity),

(v) graphical themes,

(vi) authentication methods (can require username and password accessibility),

(vii) enrollment methods,

(viii) content filters.

Integrating a CMS with the LMS—both web applications—makes a single sign-on solution for the u-learning community realistic. Another web application yet to be integrated is the Elgg social networking platform with support for blogs, wikis, communities, and other things that could be associated with the CMS.

Elgg provides the necessary functionality to allow you to run your own social networking site, whether publicly (like Facebook) or internally on a networked intranet (like Microsoft SharePoint) [29].

5. Discussions and Its Significance

In a conventional e-learning environment, students learn and practice in the cyber world wherein complex operations or problem-solving procedures are usually trained in a web-based learning environment that simulates the scenarios of the problem domain. These learning or training approachs are helpful to the learners in identifying the problem to be coped with, but it could be difficult for the students to learn the problem-solving skills without observing and practicing in the real world.

This paper of an efficient and interactive WSN-based ubiquitous learning model proposes a web-based interactive social learning model based on wireless sensor networks which are able to sense the personal and environmental contexts to provide adaptive supports to the learners.

An adaptive model of technology selection is a requirement to ensure that the needs of the learners are met. The interactive u-learning is a learning approach which asserts that knowledge and the learning of knowledge are distributive and are not located in any given place but rather consist of the network of connections formed from experience and interactions with a knowing community, and the digital-age learner is thinking and interacting in new ways (Figure 10).

The proposed WSN-Based u-learning model helps produce positive outcomes as they are delivered and evaluated, supported by holistic approaches that include appropriate policies, infrastructure, professional development, and curricula. It is also indicated that a more technology-rich environment delivers greater impacts.

For students, this model can provide an educationally-superior alternative to traditional lectures, in which learning can take place outside of the classrooms. It can also provide a model for students on how to become self-directed, independent learners, which may assist them to become "lifelong learners."

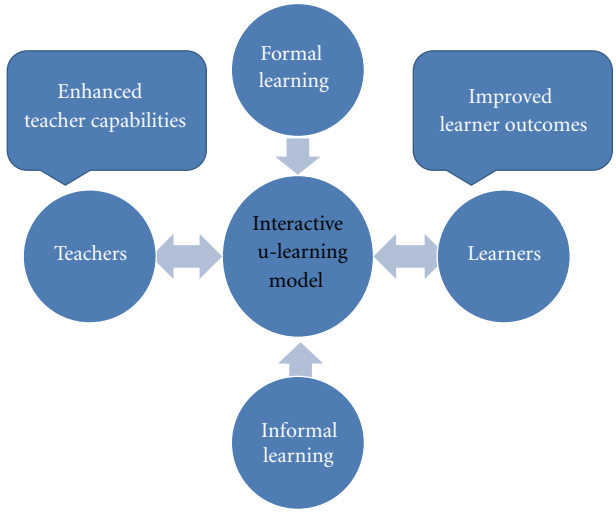


Figure 10: u-learning community significance.

The u-learning model can help increase student engagement, motivation, and attendance which are the key requisites for learning. It could effectively improve performance on core subjects and foster the development of 21st century skills. Thus, it has the potential of creating a sense of engagement, motivation, retention, progression, excitement, and involvement which could promote new forms of attainment which were not adequately measured by traditional assessment.

For teachers, this model provides them with the opportunity to test students in real business situations and new methods to evaluate each student's learning. These may cause changes in their work patterns and even change the teacher's entire approach to teaching and learning. It is not just a change of knowledge and content but the pedagogy and relationships between teachers and learners as well.

The u-learning model can improve communications between students, and between students and teachers, thus, it promotes interactivity, rapidity of response, and objectivity of marking, enabling personalized support, learner-led pacing, new forms of access to learning, and so forth. With this technology, teachers can access tools that enable them to deliver customized assessments and gain immediate feedback on individuals and class progress. With this feedback, they can provide learning opportunities using remediation and enrichment to deliver more differentiated instructions that meet each learner's needs, and thus, promote student-centered teaching and preparation.

The proposed model emphasized that the u-learning model is much more organized than traditional and conventional one-to-one instruction with its ability to interface learners with the real-world scenario, and such a systematic learning facility could increase learning efficiency. It can also indicate that students would prefer to learn with such a u-learning system because of its convenience and interactive and innovative system.

In addition, the u-learning model proposed in this study can effectively and efficiently improve the performance of training complex problem-solving skills in the real world. This model is highly recommended as an innovative approach to learning science environments, such as specific physics or chemistry experiments.

6. Conclusion

The ubiquitous learning concept has gone beyond portable computers. Technological innovations will be embedded and blended into our daily chores as new technology evolves and more pervasive forms of technology emerge. We tend to easily adapt to the use of technologies and pedagogies that emerge in this age of progress and great change.

The advantages of an adaptive learning environment combine with the benefits of ubiquitous computing and the flexibility of mobile devices through the development of ubiquitous learning environment. Students have the freedom to learn within a learning environment which offers adaptability to their individual needs and learning styles, as well as the flexibility of pervasive and unobtrusive computer systems.

u-learning is a learning paradigm which takes place in a ubiquitous computing environment that enables learning the right thing at the right place and time in the right way. It is an expansion of previous learning paradigms as we move from conventional learning to electronic learning to mobile learning and now to ubiquitous learning. This u-learning model is guided by the theory of connectivism which asserts that knowledge and the learning of knowledge are distributive and are not located in any given place but rather consist of the network of connections formed from experiences and interactions with a knowing community.

u-learning model is designed as an interactive social learning model based on wireless sensor networks that aim at accommodating learners in their learning style by providing adequate information at anytime and anywhere as they wish for it. Hence, u-learning model is web-based and incorporates e-learning, m-learning, wireless sensor networks, and the ubiquitous environment to provide mobility and flexibility to learners as well as the teachers.

In light of these developments, educators must update learning and ensure that their pedagogical theories are aligned with the active and collaborative nature of such technologies. They must consider the opportunities it provides for the delivery of enhanced learning.

Acknowledgment

This paper was supported by the Security Engineering Research Center, granted by the Korea Ministry of Knowledge Economy.

References

[1] H. C. Chu, G. J. Hwang, S. X. Huang, and T. T. Wu, "A knowledge engineering approach to developing e-libraries for mobile learning," *Electronic Library*, vol. 26, no. 3, pp. 303–317, 2008.

[2] G. J. Hwang, "A conceptual map model for developing intelligent tutoring systems," *Computers and Education*, vol. 40, no. 3, pp. 217–235, 2003.

[3] G. H. Hwang, J. M. Chen, G. J. Hwang, and H. C. Chu, "A time scale-oriented approach for building medical expert systems," *Expert Systems with Applications*, vol. 31, no. 2, pp. 299–308, 2006.

[4] G. J. Hwang, C. C. Tsai, and S. J. H. Yang, "Criteria, strategies and research issues of context-aware ubiquitous learning," *Educational Technology and Society*, vol. 11, no. 2, pp. 81–91, 2008.

[5] L. M. Feeney, B. Ahlgren, and A. Westerlund, "Spontaneous networking: an application-oriented approach to ad hoc networking," *IEEE Communications Magazine*, vol. 39, no. 6, pp. 176–181, 2001.

[6] T. Kindberg and A. Fox, "System software for ubiquitous computing," *IEEE Pervasive Computing*, vol. 1, no. 1, pp. 70–81, 2002.

[7] G. D. Abowd, "An experiment with the instrumentation of a living educational environment," *IBM Systems Journal*, vol. 38, no. 4, pp. 508–530, 1999.

[8] T. Uemukai, T. Hara, and S. Nishio, "A method for selecting output data from ubiquitous terminals in a ubiquitous computing environment," in *Proceedings of the 24th International Conference on Distributed Computing Systems Workshops (ICDCSW '04)*, pp. 562–567, Tokyo, Japan, 2004.

[9] O. Kwon, K. Yoo, and E. Suh, "ubiES: an intelligent expert system for proactive services deploying ubiquitous computing technologies," in *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, p. 85, Hawaii, Hawaii, USA, January 2005.

[10] S. K. Mostefaoui, A. T. Bouzid, and B. Hirsbrunner, "Using context information for service discovery and composition," in *Proceedings of the 5th International Conference on Information Integration and Web-Based Applications and Services (IIWAS '03)*, Jakarta, Indonesia, 2003.

[11] Y. Isoda, S. Kurakake, and H. Nakano, "Ubiquitous sensors based human behavior modeling and recognition using a spatio-temporal representation of user states," in *Proceedings of the 18th International Conference on Advanced Information Networking and Application*, vol. 1, pp. 512–517, Fukuoka, Japan, 2004.

[12] H. Ogata, "Features of ubiquitous learning from computer supported ubiquitous learning environment for language learning," *Journal Transactions Processing Society of Japan*, vol. 45, no. 10, pp. 2354–2363, 2003.

[13] M. Collins, *Adult Education as Vocation: A Critical Role for the Adult Educator*, Routledge, New York, NY, USA, 1991.

[14] University of Illinois College of Education Ubiquitious Learning Institute, http://www.ed.uiuc.edu/uli/.

[15] T. Stockwell, *Accelerated Learning in Theory and Practice*, European Foundation for Education, Communication and Teaching, 1998.

[16] S. J. H. Yang, B. C. W. Lan, B. J. D. Wu, and A. C. N. Chang, "Context aware service oriented architecture for web based learning," *International Journal of Advance learning Technology*, vol. 2, no. 4, 2005.

[17] W. Carine, B. Loris, P. Sylvie, and B. Nicolas, "The Baghera project: a multi-agent architecture for human learning," http://julita.usask.ca/mable/webber.pdf.

[18] G. Zhang, Q. Jin, and M. Lin, "A framework of social interaction support for ubiquitous learning," in *Proceedings of the19th International Conference on Advanced Information Networking and Applications*, pp. 639–643, March 2005.

[19] C. Townsend and S. Arms, *Wireless Sensor Networks: Principles and Applications*, Chapter 22, MicroStrain, 2005.

[20] Y. K. Wang, "Context awareness and adaptation in mobile learning," in *Proceedings of the 2nd IEEE International Workshop on Wireless and Mobile Technologies in Education*, pp. 154–158, March 2004.

[21] http://www.webopedia.com/TERM/M/Mobile_IP.html.

[22] http://en.wikipedia.org/wiki/Mobile_IP.

[23] S. Downes, Coonectivism and Connective Knowledge, 2011, http://cck11.mooc.ca/week1.htm.

[24] G. Siemens, Connectivism: A Learning Theory for the Digital Age, http://www.connectivism.ca/ and http://www.itdl.org/journal/jan_05/article01.htm.

[25] S. Yahya, E. A. Ahmad, and K. Abd Jalil, "The definition and characteristics of ubiquitous learning: a discussion," *International Journal of Education and Development Using Information and Communication Technology*, vol. 6, no. 1, 2010.

[26] G.-J. Hwang, T.-C. Yang, C.-C. Tsai, and S. J. H. Yang, "A context-aware ubiquitous learning environment for conducting complex science experiments," *Computers and Education*, vol. 53, no. 2, pp. 402–413, 2009.

[27] http://www.joomla.org/.

[28] http://docs.moodle.org/.

[29] http://docs.elgg.org/wiki/Main_Page.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 409048, 7 pages
doi:10.1155/2012/409048

Research Article

Real-Time Train Wheel Condition Monitoring by Fiber Bragg Grating Sensors

Chuliang Wei,[1] Qin Xin,[2] W. H. Chung,[3] Shun-yee Liu,[3] Hwa-yaw Tam,[3] and S. L. Ho[3]

[1] *Department of Electronic Engineering, Shantou University, Guangdong 515063, China*
[2] *Department of Applied Mathematics, Université catholique de Louvain 1348, Louvain-la-Neuve, Belgium*
[3] *Department of Electrical Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong*

Correspondence should be addressed to Qin Xin, xin@simula.no

Received 25 July 2011; Accepted 8 August 2011

Academic Editor: Tai Hoon Kim

Wheel defects on trains, such as flat wheels and out-of-roundness, inevitably jeopardize the safety of railway operations. Regular visual inspection and checking by experienced workers are the commonly adopted practice to identify wheel defects. However, the defects may not be spotted in time. Therefore, an automatic, remote-sensing, reliable, and accurate monitoring system for wheel condition is always desirable. The paper describes a real-time system to monitor wheel defects based on fiber Bragg grating sensors. Track strain response upon wheel-rail interaction is measured and processed to generate a condition index which directly reflects the wheel condition. This approach is verified by extensive field test, and the preliminary results show that this electromagnetic-immune system provides an effective alternative for wheel defects detection. The system significantly increases the efficiency of maintenance management and reduces the cost for defects detection, and more importantly, avoids derailment timely.

1. Introduction

Condition monitoring measures are crucial to ensure safe and cost-effective train operation in the railroad transportation industry. A well-designed monitoring system substantially reduces hardware maintenance cost and improves service quality and overall safety. Wheel condition monitoring is one of the critical features in railway condition monitoring system. Typical train wheel defects include flat wheels, out-of-roundness (OOR), spalling, and shelling [1–3]. If wheel defects are not detected and rectified in time, the wheel may deteriorate rapidly and induce more frictional force that inflicts further defect on both wheels and rails.

Nowadays, railway operators usually detect wheel defects by visual checking with experienced workers on a regular basis. Passenger complaint or driver report of excessive vibration is another means of identifying wheel defects. Moreover, periodical scheduled wheel reprofiling according to engineering experiences without defects identification is also employed. These methods are useful in general but they do not guarantee in-time identification of wheel defects. The railway operators have to bear the high cost for regular wheel-checking and reprofiling, and the comfort of passengers is adversely affected due to high vibration. Worst still, there is always a higher risk of derailment. Thus, a real-time monitoring system which is reliable, safety-proven, and cost-effective for handling wheel defects problem is urgently required.

Many studies have been carried out to realize wheel defect detection and most of them are based on the analysis of wheel-rail interaction [2–6]. Attivissimo et al. used a laser diode and a CCD camera to measure wheel and rail head profiles and hence to evaluate the wheel-rail interaction quality [2]. The system is able to discover anomalies of wheel-rail contact, but not resolving the defects of wheel or rail. In addition, the laser source and camera require precision setup which is difficult for practical railroad application. To study possible causes and effects of OOR wheels, strain gauges and accelerometers were employed to measure the vertical wheel-rail contacting force and track response [3, 4], a number of 99 selected wheels were tested over a 100,000 km traveling distance. Since the sensitivity of the system is relatively low, only the wheels with local defects of 0.5 m in length can be identified. Sensors based

on ultrasonic [5] and acoustic [7] techniques have also been employed to measure the wheel and rail conditions, but the performance of these sensors are easily compromised under electromagnetic interference (EMI) railroad environment. Besides, EMI free fiber optics sensor [8] has been used to study the wheel defect problem, speckle interferometry in multimode optical fiber was developed to detect flat wheels of train. However, other types of defects such as local spall or polygonal wheel cannot be retrieved.

In this paper, we propose a real-time wheel-defect detection system based on fiber Bragg grating (FBG) sensors. The sensors measure the rail strain response upon wheel-rail interaction and the frequency component that solely reveals the quality of the interaction are extracted from the signal and processed in order to deduce the defects of passing wheels. One of the advantages of this sensor system is that both the sensors and connecting fibers installed at the railroad side are passive to EMI and they require no electric power until the head-end measurement equipment which could be tens of kilometers away from the measurement points. This feature is particularly favorable to the modern electrified railway system since the sensing network is immunized from EMI. In addition, the system allows in-service and real-time monitoring of wheel condition, which is attractive to the railway industry. Moreover, the system can also be integrated with other railroad FBG sensing applications [9–13] that have been reported previously.

The organization of this paper is as follows. The basic concept of FBG sensor and its characteristics are first reviewed, followed by the descriptions of the sensor packaging and field test; the measurement results and the algorithm to generate the condition index (CI) of wheel are presented; discussions and conclusions are also given at the subsequent sections.

2. FBG Sensor and Wavelength Interrogation System

FBG sensor is an in-fiber narrowband reflective optical filter resulted from periodical variation of refractive index inside the core of optical fiber. A conventional way to fabricate a FBG sensor is to illuminate a short section of optical fiber by ultraviolet laser source under a phase mask, the periodical pattern of the mask is photo-imprinted onto the optical fiber by modulating the refractive index of the fiber core. The Bragg wavelength λ_B of FBG sensor written in single-mode optical fiber is given below [14]:

$$\lambda_B = 2n\Lambda, \tag{1}$$

where n is the effective refractive index of the fiber core and Λ is the period of the refractive index modulation. Figures 1(a) and 1(b) illustrate an FBG sensor and its spectral characteristics. When a broadband light source is passed onto a sensor, a narrowband optical spectrum centered at λ_B is reflected while others pass through. Mechanical and thermal perturbations alter the modulation pitch as well as refractive index of the FBG and hence λ_B. By measuring the wavelength change of the FBG sensor, these perturbations can be determined. Typical wavelength changes of FBG written in standard single-mode fiber at 1550 nm region due to mechanical strain and temperature variation are ~1 pm/$\mu\varepsilon$ and ~11 pm/°C, respectively. Because of its compact size, low optical loss, self-referencing, and wavelength multiplexing capability, FBG sensors have been widely applied in a wide range of condition monitoring applications [9, 15–19].

An interrogator is employed to measure the wavelength of FBG sensor array. Figure 1(c) shows a simplified schematic of a typical FBG interrogator. To produce a wavelength tunable optical source with high output power for accurate wavelength measurement, scanning ring laser consists of an optical amplifier and a tunable narrowband Fabry-Perot filter in a ring cavity configuration is employed, the laser scans the FBG sensor array periodically and the signal reflected from the FBG sensors is directed to the photodetector through an optical circulator. The photodetector performs optical to electrical domain conversion, any peak in the converted electric signal reflects the existence of an FBG and the wavelength of the FBG can be retrieved by the timing of scanning signal. Commercially available interrogators usually have four measurement channels with a wavelength measuring range of ~80 nm, depending on the spectral occupancy of the FBG sensors in the array. Up to 40 serially connected FBG sensors (assuming 1 nm measuring range with 0.5 nm guard-band at both longer and shorter wavelength sides) can be measured simultaneously at a maximum sampling rate of 2 kHz by a single measurement channel. In this study, a four-channel interrogator with a sampling rate of 1 kHz is used.

3. Sensor Packaging and Field Test

Phase masks and ultraviolet excimer laser are used to fabricate the 10 mm-long FBG sensors in this study. To improve the photosensitivity of the standard single-mode fiber for FBG inscription, the fiber is kept in high pressure hydrogen vessel for a week. After FBG sensor fabrication, the fiber is annealed in temperature chamber and the FBG sensor section is recoated with the acrylate material. A 1 mm thickness stainless steel sheet with a dimension of 50 mm (length) and 12 mm (width) is machined using photochemical etching technique to form the sensor package as shown in Figure 2(a). This sensor package is designed in such a way that it allows a good mechanical contact between the FBG sensor and the rail surface while providing sufficient protection to the fiber. A 10 mm (length) and 8 mm (width) rectangular-shape void at the center of the package is used to place the FBG sensor during assembling, four elliptical recessed pads with a depth of 0.7 mm are used for electric spot-welding to the rail during on-site sensor installation. U-shape recessed grooves with a width and depth of 1 mm and 0.5 mm, respectively, are designed for fiber alignment and attachment purposes.

The sensor package is tested substantially by a calibration platform shown in Figure 2(b), a stainless steel triangular beam with one end fixed and the other end coupled to a mechanical joint which converts the rotational motion generated by a motor to a bending movement of the beam.

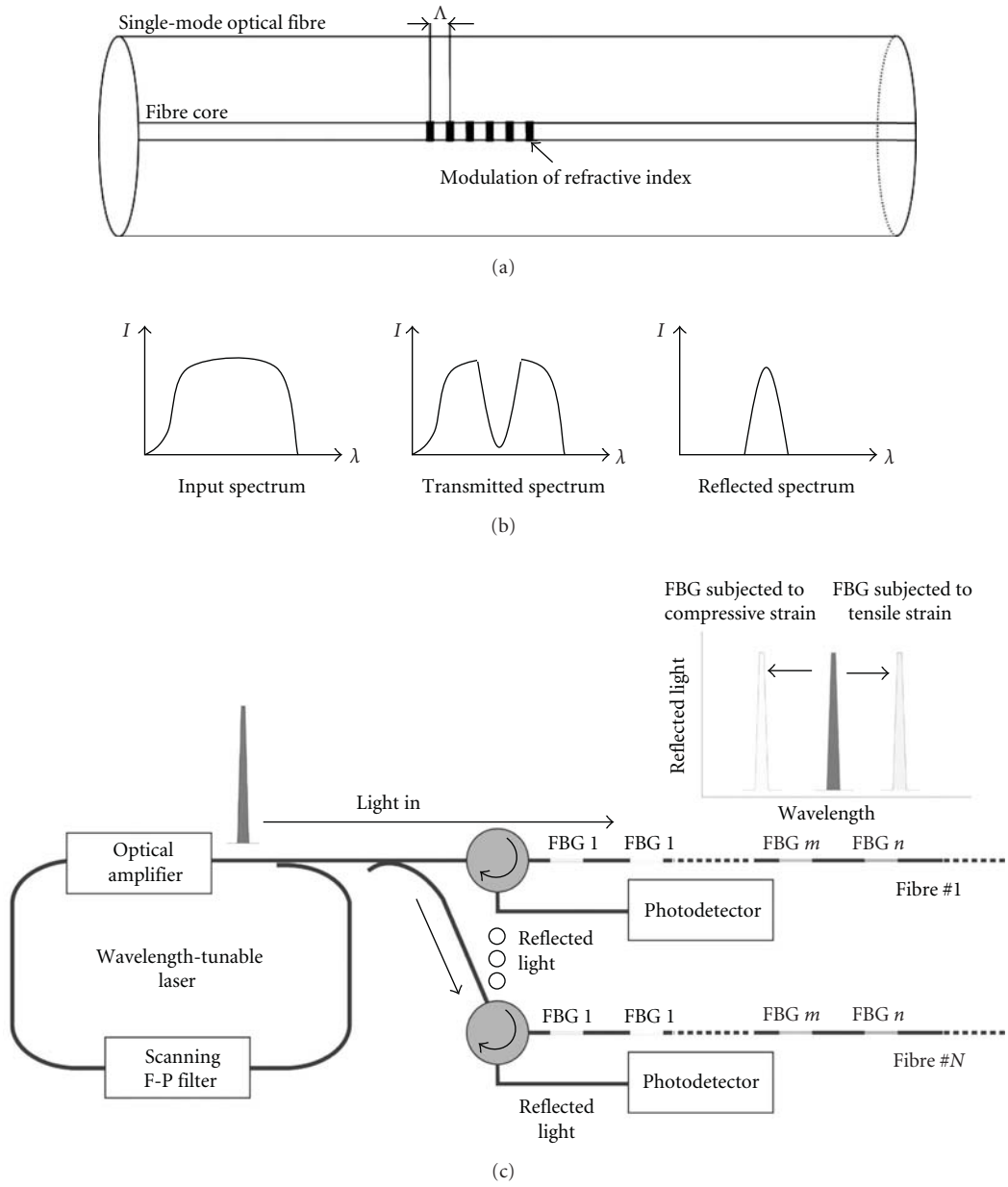
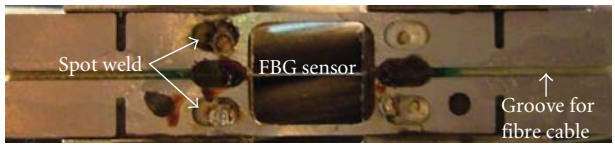


Figure 1: (a) Illustration of an FBG sensor; (b) the spectral response of FBG sensor; (c) block diagram of an FBG interrogator.

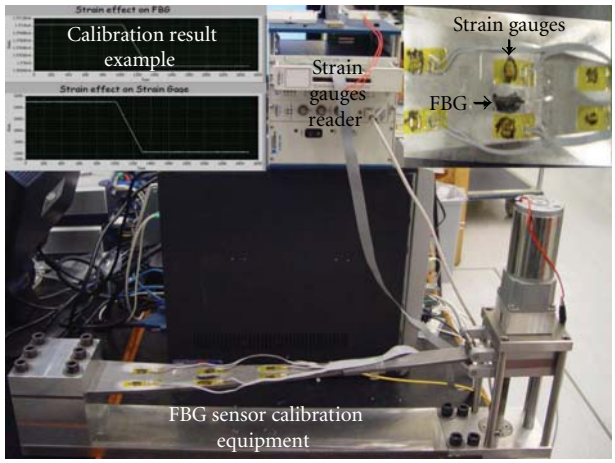
Strain gauges and the packaged FBG sensors are attached onto the beam for a back-to-back comparison. The motor, strain gauges reader, and the FBG interrogator are controlled by a computer for data logging. Performance and robustness of the packaged FBG sensors are maintained after thousands of bending cycles.

The field test has been performed in the East Rail line of the Hong Kong Mass Transit Railway (MTR). It is a heavily used suburban railway line with both passenger and freight services connecting Hong Kong urban districts to the border with mainland China. There are 14 intermediate stations and its length is over 30 km. The majority of track-form is the conventional ballast track with a maximum train speed of 130 km/h. Two independent pairs of tracks are used to provide traveling services for two opposite (i.e., north-bound and south-bound) directions.

Four sensors have been installed by spot-welding onto two pairs of rail tracks near the Ho Tung Lau (HTL) depot in the East Rail line. Train speed at the measurement point is 50 to 90 km/h with ballast track-form. The sensors are installed near the foot of the rail which is the most sensitive and practical position [13] for measuring the wheel-rail interaction. The rail has been slightly polished to provide a clean contact surface area for spot welding of the sensors. Arrangement of the sensor installation is shown in Figure 3, FBGs 1 and 2 are used to measure north-direction track; and FBGs 3 and 4 are for south-direction track. A pair of sensors (FBGs 1 and 2 or FBGs 3 and 4) is installed at the



(a)



(b)

Figure 2: (a) Package of FBG sensor; (b) calibration platform in laboratory.

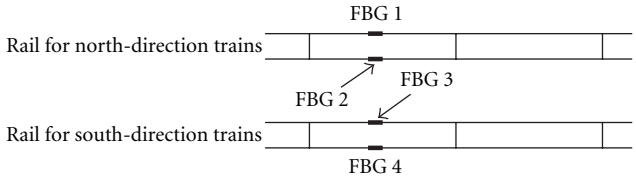


Figure 3: The FBG sensors installation scheme.

same location to ensure the two wheels of an axle passing the two measurement points at the same instant. The four sensors are connected in series as an array and linked to the interrogator through 2 km of outdoor optical cables. Real-time cameras and RFID receivers are also installed at two ends of the measurement area for train identification and mapping of sensor measurement results. The interrogator and other measurement equipments are operated in the control center of the HTL depot.

4. Results and Discussions

Figure 4(a) shows the measured strain change of an FBG sensor when a typical 12-car passenger train passes a measurement point. The maximum strain change is ~250 $\mu\varepsilon$ (i.e., the wavelength change is ~0.25 nm). The 48 axles of the train are clearly indicated in the signal, as illustrated by the 48 peaks. The characteristics of strain change experienced by the rail induced by a passing train depend on the speed and weight of the train, as well as conditions of rail and wheels. The low-frequency component of the signal mainly comes from the axle load of the train as depicted by the strain peaks of Figure 4(a), those peaks also show the slightly heavy motor cars, and this signal is essential information for studying train axle counting system [13].

Figures 4(b) to 4(d) show the enlarged sections of Figure 4(a) that indicate the measured strain responses of the first to third cars of the train accordingly. Because these results are measured by the same sensor within a short period of time, they are therefore virtually collected under the same rail condition and reflected only the difference in the condition of the wheels. Obviously, Figure 4(b) shows a smooth strain signal generated by the first car while the signal generated by the third car as shown in Figure 4(d) is the noisiest among the three. Table 1 shows the wheel condition information of the three cars which is provided by the MTR, the OOR measurement in this table is conducted by using contact gauge in maintenance depot. The OOR of the wheels of the third car is much higher than those of the other two ones, and there is some thread wear in the wheels of the second car. In other words, the wheel condition of the first car is the best and that of the third car is the worst. Imperfect wheels of the third car exert periodic impact force onto the track and hence induce the measured uneven strain impulse as depicted by the noisy signal of Figure 4(d). Therefore smoother strain signal represents better condition of the wheels and this is consistent with the studies described previously [3, 6].

A condition index (CI) system has been developed to express the wheel condition quantitatively. This system processes the strain signal of the four wheels (left or right side) of a train car and generates a CI that essentially indicates the condition of the four wheels (if the condition of any wheel is not in good condition, the system can generate a CI to indicate the wheels of the train car need to be maintained, but cannot distinguish which wheel is not in good condition because the system processes the strain signal of the four wheels as a whole. The reason of processing four wheels as a whole is that railway operators normally do wheel reprofiling on a train car basis). The low-frequency component of the strain signal is first filtered out by using a low-pass filter with a cutoff frequency of 20 Hz, Figure 5(a) gives the low-pass filter result of the signal of Figure 4(d). This signal basically consists of the information of axle loading of the train induced by its weight. The train speed (v) can be deduced from the distance between two wheels and the time between the peaks of the signal. Figure 5(b) shows the high-pass-filtered signal which carries the wheel condition information, and the signal is then decomposed to different frequency components by fast Fourier transform (FFT), as shown in Figure 5(c). The distribution shows that the majority of the strain change is located in the frequency range of 20 to 200 Hz while others are mainly noise at the frequency above 200 Hz. Further experimental works has confirmed that higher strain change over this frequency range represents a worse wheel or rail conditions. Moreover, higher train speed (v) generates higher average strain change which increases 0.01 $\mu\varepsilon$ for a 1 km/h increment with the train speed running at 50 to 90 km/h. The equation for calculating

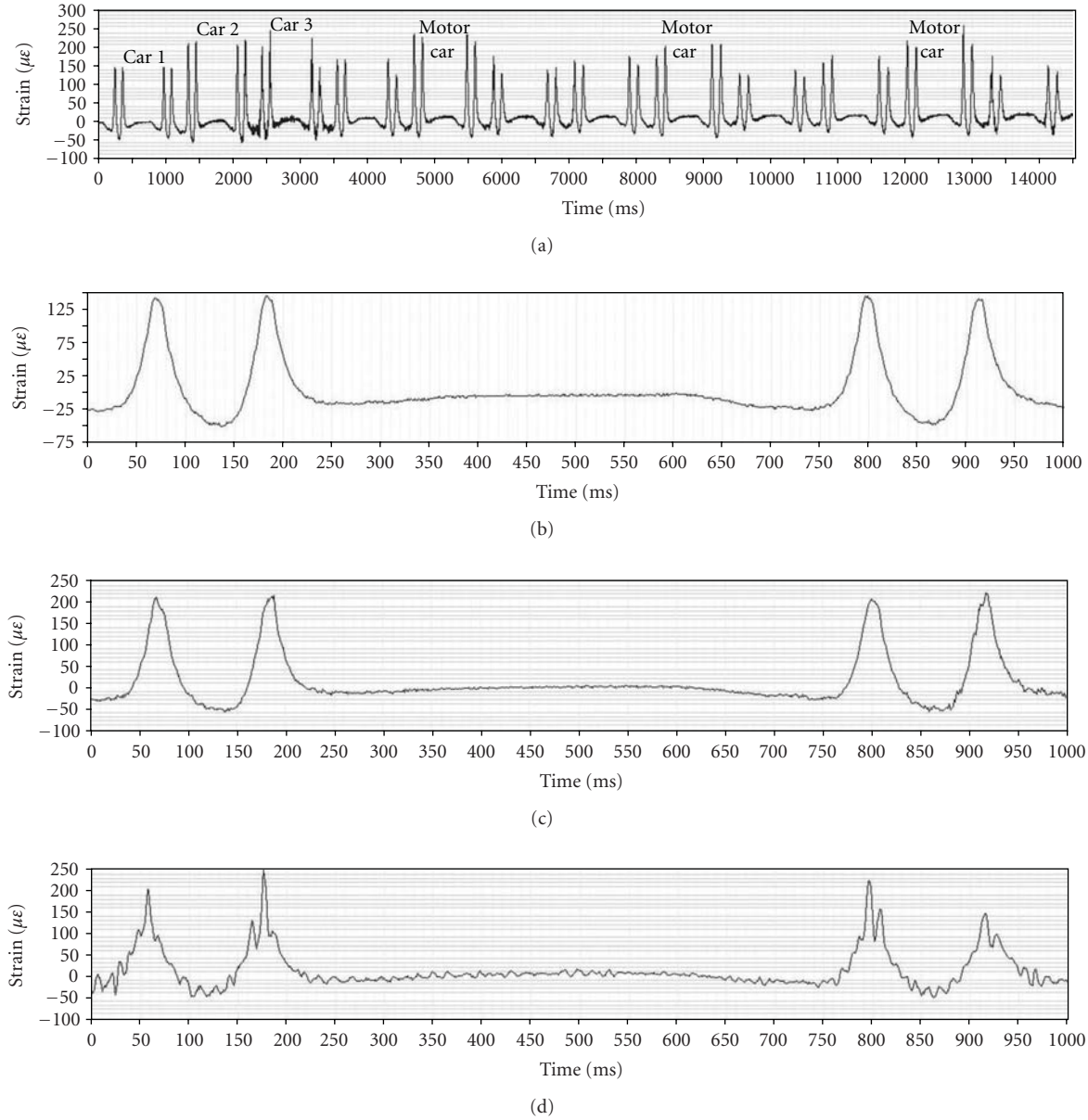


Figure 4: The measured strain change of (a) a 12-car passage train; (b) the first car; (c) the second car; (d) the third car.

Table 1: Wheel condition information of the three train cars.

Car number	OOR (on average)	Flat	cavity	Flange thin	Ellipse	Tread wear	Vibration
1	0.36	\	\	\	\	\	\
2	0.38	\	\	\	\	*	\
3	0.65	\	\	\	\	\	\

CI is therefore given by

$$\mathrm{CI} = \frac{\bar{\varepsilon}}{v} \times A, \tag{2}$$

where $\bar{\varepsilon}$ is the average value of the strain changes in the frequency range of 20 to 200 Hz, A is a scaling factor requires for matching the CI value within a range of 0 to 10.

The FBG monitoring system with CI was tested continuously over ten months, and the CIs of the twenty-nine passenger trains running in East Rail have been collected. Figures 6(a) and 6(b) show the evolutions of the CI over the 10-month test period of train cars 027_039 and 072_049, respectively, a single dot represents the train car passing

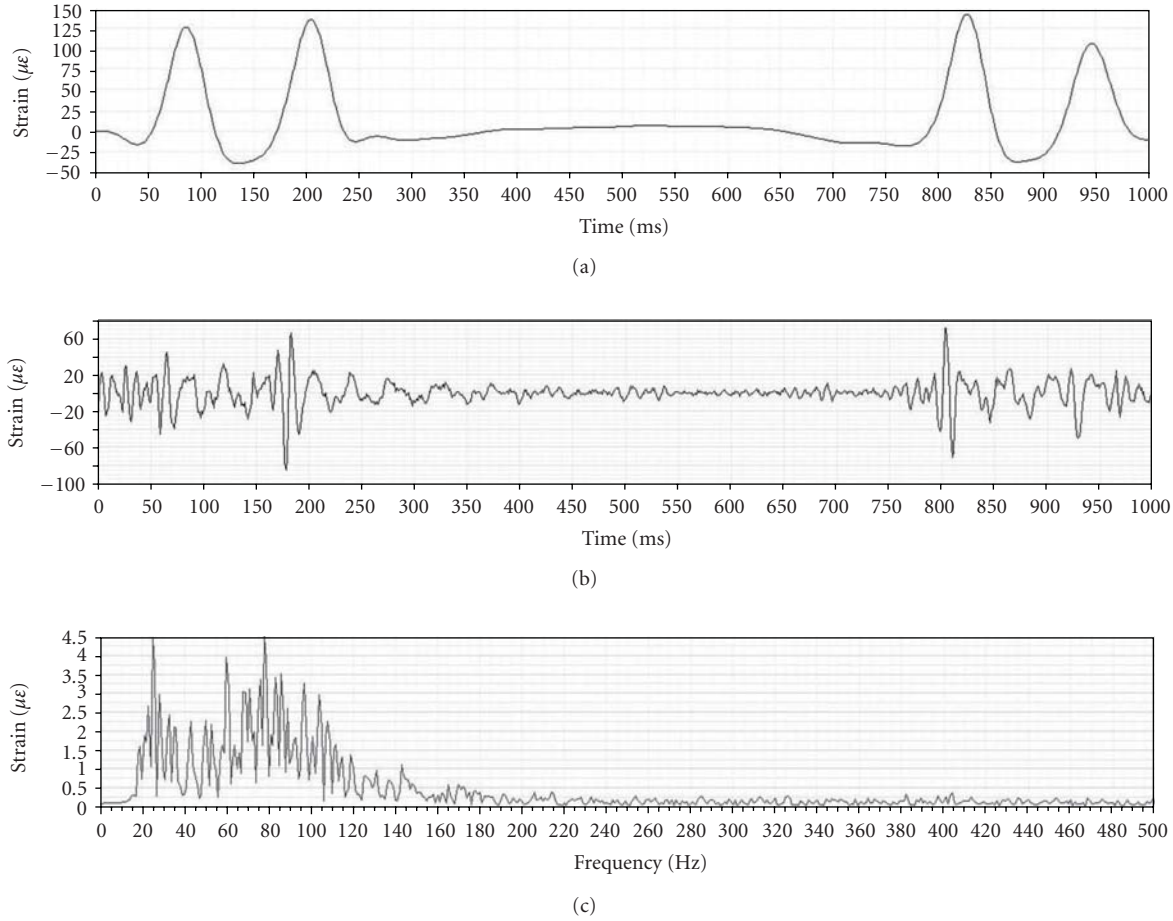


Figure 5: (a) The low-passed-filtered result of the signal in Figure 4(d); (b) the high-passed-filtered result of the signal in Figure 4(d); (c) frequency spectrum of the signal in (b) processed by FFT.

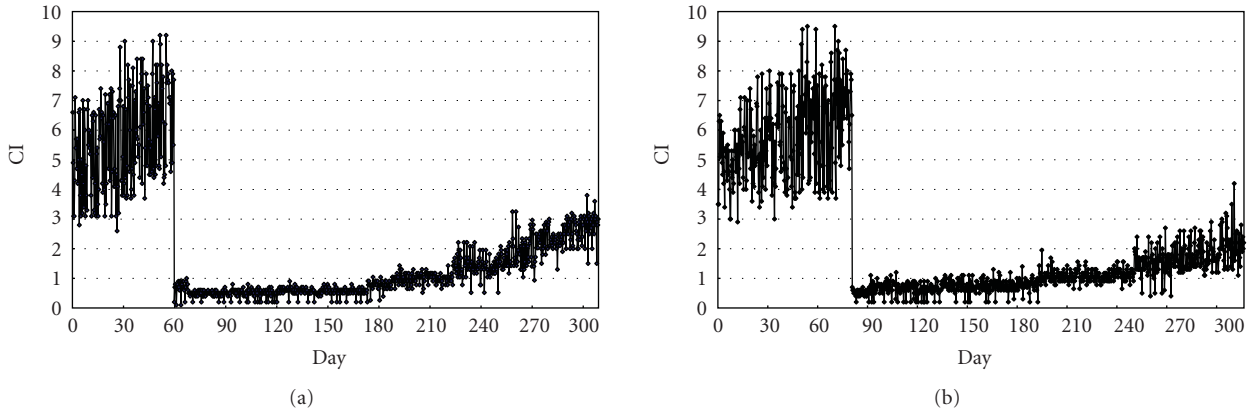


Figure 6: CI generated over ten months' period for (a) the left-side wheel of the train car 027_039; (b) the right-side wheel of the train car 072_049.

the measurement point once. It can be noted that the CI gradually increases for the first two/three months during the test until the wheels are reprofiled by depot grinding machine, the CI is substantially reduced from an average value of ∼7 to less than ∼1. After wheel reprofiling, the CI increases slightly and deteriorates again after ∼4 months in operation. This can be explained by the fact that newly turned wheels produce very little uneven strain on the rail and once a defect formed on a single wheel, deterioration of this wheel is accelerated and this also affects the other wheels

of the same bogie and hence degrades the overall wheel condition of the train car measured by the sensor system. It is also interesting to note from the figures that under normal train operation condition, the CI takes about 12 months to rise from ~1 of newly turned wheels to ~7 of just before wheels reprofiling, this time frame matches very well with the regular wheels reprofiling schedule exercised by the railway operator.

5. Conclusions

A real-time train wheel condition monitoring system using FBG sensors to measure the track strain response upon wheel-rail interaction is proposed and developed. The principle of FBG sensor and its optical characteristics are briefly reviewed, together with the description of the sensor packaging and field test. An algorithm to generate the CI value for indicating the train wheel condition is described, followed by the field test results and discussions.

The system has been proven to be able to continuously monitor the conditions of the wheel and identify wheel defects in time. The system has the advantage of providing real-time and in-service measurement of wheel condition. Moreover, FBG sensors used in this system are very low in optical loss and immunity from EMI is particular suitable for electrified railways with track spanning over tens or even hundreds of kilometers. Further test and verifications with wheel and rail condition data are still required to ensure the long-term reliability of the proposed system. To improve the measurement accuracy and overall integrity of the system, a more intelligent analysis technique based on the previous wheel and rail conditions information and the results of a measurement point with more FBG sensors installed is now under investigation. A more advanced FBG-based system which can differentiate the defects of wheel and rail, as well as their types and levels, is the next stage of the real-time railway condition monitoring development.

Acknowledgments

The authors would like to express their gratitude to financial support from Shantou University and the Hong Kong Polytechnic University, and the technical support from the Hong Kong MTR.

References

[1] J. C. O. Nielsen and A. Johansson, "Out-of-round railway wheels-a literature survey," *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 214, no. 2, pp. 79–91, 2000.

[2] F. Attivissimo, A. Danese, N. Giaquinto, and P. Sforza, "A railway measurement system to evaluate the wheel-rail interaction quality," *IEEE Transactions on Instrumentation and Measurement*, vol. 56, no. 5, pp. 1583–1589, 2007.

[3] A. Johansson and J. C. O. Nielsen, "Out-of-round railway wheels—wheel-rail contact forces and track response derived from field tests and numerical simulations," *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 217, no. 2, pp. 135–145, 2003.

[4] A. Johansson, "Out-of-round railway wheels-assessment of wheel tread irregularities in train traffic," *Journal of Sound and Vibration*, vol. 293, no. 3–5, pp. 795–806, 2006.

[5] M. Guagliano and M. Pau, "An experimental-numerical approach for the analysis of internally cracked railway wheels," *Wear*, vol. 265, no. 9-10, pp. 1387–1395, 2008.

[6] P. Remington and J. Webb, "Estimation of wheel/rail interaction forces in the contact area due to roughness," *Journal of Sound and Vibration*, vol. 193, no. 1, pp. 83–102, 1996.

[7] E. Matzan, "System for detection of defects in railroad car wheels," U.S. Patent 7213789, 2007.

[8] D. R. Anderson, "Detecting flat wheels with a fiber-optic sensor," in *Proceedings of the ASME/IEEE 2006 Joint Rail Conference (JRC '06)*, pp. 25–30, April 2006.

[9] H. Y. Tam, S. Y. Liu, B. O. Guan, W. H. Chung, T. H. T. Chan, and L. K. Cheng, "Fiber Bragg grating sensors for structural and railway applications," in *Advanced Sensor Systems and Applications II*, vol. 5634 of *Proceedings of SPIE*, Beijing, China, November 2004.

[10] K. Y. Lee, K. K. Lee, and S. L. Ho, "Exploration of using FBG sensor for derailment detector," *WSEAS Transactions Topics Systems*, vol. 3, pp. 2433–2439, 2004.

[11] K. Y. Lee, K. K. Lee, and S. L. Ho, "Exploration of using FBG sensor for axle counter in railway engineering," *WSEAS Transactions Topics Systems*, vol. 3, pp. 2440–2447, 2004.

[12] H. Y. Tam, S. Y. Liu, W. H. Chung, T. K. Ho, S. L. Ho, and K. K. Lee, "Smart railway sensor network using fiber Bragg grating sensors as sensing elements," in *Smart Materials and Smart Structures Technology*, vol. 1, pp. 159–169, Techno-Press, Daejeon, Korea, 2005.

[13] C. L. Wei, C. C. Lai, S. Y. Liu et al., "A fiber Bragg grating sensor system for train axle counting," *IEEE Sensors Journal*, vol. 10, no. 12, pp. 1905–1912, 2010.

[14] A. D. Kersey, M. A. Davis, H. J. Patrick et al., "Fiber grating sensors," *Journal of Lightwave Technology*, vol. 15, no. 8, pp. 1442–1462, 1997.

[15] T. H. T. Chan, L. Yu, H. Y. Tam et al., "Fiber Bragg grating sensors for structural health monitoring of Tsing Ma bridge: background and experimental observation," *Engineering Structures*, vol. 28, no. 5, pp. 648–659, 2006.

[16] E. J. Friebele, C. G. Askins, A. B. Bosse et al., "Optical fiber sensors for spacecraft applications," *Smart Materials and Structures*, vol. 8, no. 6, pp. 813–838, 1999.

[17] N. E. Fisher, J. Surowiec, D. J. Webb et al., "In-fibre Bragg gratings for ultrasonic medical applications," *Measurement Science and Technology*, vol. 8, no. 10, pp. 1050–1054, 1997.

[18] B. O. Guan, H. Y. Tam, and S. Y. Liu, "Temperature-independent fiber Bragg grating tilt sensor," *IEEE Photonics Technology Letters*, vol. 16, no. 1, pp. 224–226, 2004.

[19] Y. J. Rao, "Recent progress in applications of in-fibre Bragg grating sensors," *Optics and Lasers in Engineering*, vol. 31, no. 4, pp. 297–324, 1999.