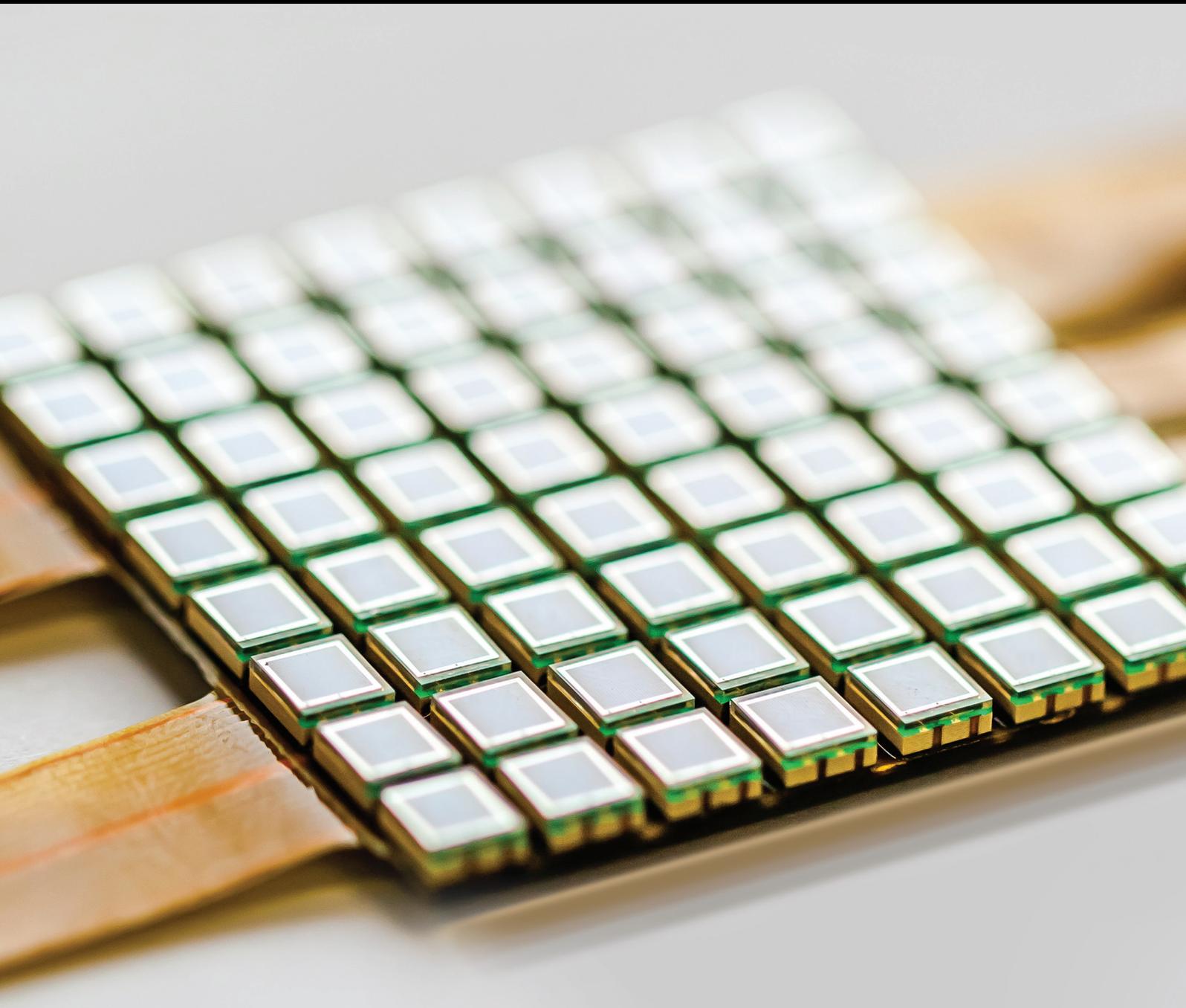


# Embedded Systems for Mobile Sensors

Guest Editors: Marco Anisetti, Valerio Bellandi, Abdellah Chehri, Yurong Qian, and Gwanggil Jeon





---

# **Embedded Systems for Mobile Sensors**

Journal of Sensors

---

## **Embedded Systems for Mobile Sensors**

Guest Editors: Marco Anisetti, Valerio Bellandi, Abdellah Chehri,  
Yurong Qian, and Gwanggil Jeon



---

Copyright © 2016 Hindawi Publishing Corporation. All rights reserved.

This is a special issue published in "Journal of Sensors." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Editorial Board

Harith Ahmad, Malaysia  
Francisco J. Arregui, Spain  
Francesco Baldini, Italy  
Fernando Benito-Lopez, Ireland  
Romeo Bernini, Italy  
Shekhar Bhansali, USA  
Wojtek J. Bock, Canada  
Hubert Brändle, Switzerland  
Davide Brunelli, Italy  
Paolo Bruschi, Italy  
Belén Calvo, Spain  
Stefania Campopiano, Italy  
Domenico Caputo, Italy  
Sara Casciati, Italy  
Gabriele Cazzulani, Italy  
Chi Chiu Chan, Singapore  
Nick Chaniotakis, Greece  
Nicola Cioffi, Italy  
Elisabetta Comini, Italy  
Marco Consales, Italy  
Jesus Corres, Spain  
Andrea Cusano, Italy  
Antonello Cutolo, Italy  
Dzung V. Dao, Japan  
Manel del Valle, Spain  
Ignacio Del Villar, Spain  
Francesco Dell'Olio, Italy  
Utkan Demirci, USA  
Nicola Donato, Italy  
Junhang Dong, USA  
Abdelhamid Errachid, France  
Stephane Evoy, Canada  
Vittorio Ferrari, Italy  
Luca Francioso, Italy

Laurent Francis, Belgium  
Mohammad Reza Ganjali, Iran  
Wei Gao, Japan  
Michele Giordano, Italy  
Banshi D. Gupta, India  
Clemens Heitzinger, Austria  
María del Carmen Horrillo, Spain  
Wieslaw Jakubik, Poland  
Hai-Feng Ji, USA  
Kouros Kalantar-Zadeh, Australia  
Sher Bahadar Khan, KSA  
Sang Sub Kim, Republic of Korea  
Challa Kumar, USA  
Laura M. Lechuga, Spain  
Chengkuo Lee, Singapore  
Chenzhong Li, USA  
Eduard Llobet, Spain  
Yu-Lung Lo, Taiwan  
Oleg Lupan, Moldova  
Frederick Maily, France  
Eugenio Martinelli, Italy  
Jose R. Martinez-de Dios, Spain  
Yasuko Y. Maruo, Japan  
Mike McShane, USA  
Igor L. Medintz, USA  
Fanli Meng, China  
Aldo Minardo, Italy  
Joan Ramon Morante, Spain  
Lucia Mosiello, Italy  
Masayuki Nakamura, Japan  
Calogero M. Oddo, Italy  
Marimuthu Palaniswami, Australia  
Alberto J. Palma, Spain  
Lucio Pancheri, Italy

Alain Pauly, France  
Giorgio Pennazza, Italy  
Michele Penza, Italy  
Andrea Ponzoni, Italy  
Biswajeet Pradhan, Malaysia  
Ioannis Raptis, Greece  
Armando Ricciardi, Italy  
Christos Riziotis, Greece  
Maria Luz Rodríguez-Méndez, Spain  
Albert Romano-Rodriguez, Spain  
Carlos Ruiz, Spain  
Josep Samitier, Spain  
Giorgio Sberveglieri, Italy  
Luca Schenato, Italy  
Michael J. Schöning, Germany  
Andreas Schütze, Germany  
Woosuck Shin, Japan  
Pietro Siciliano, Italy  
Vincenzo Spagnolo, Italy  
Vincenzo Stornelli, Italy  
Weilian Su, USA  
Tong Sun, UK  
Raymond Swartz, USA  
Hidekuni Takao, Japan  
Isao Takayanagi, Japan  
Guiyun Tian, UK  
Suna Timur, Turkey  
Hana Vaisocherova, Czech Republic  
Qihao Weng, USA  
Matthew J. Whelan, USA  
Stanley E. Woodard, USA  
Hai Xiao, USA

# Contents

---

## **Embedded Systems for Mobile Sensors**

Marco Anisetti, Valerio Bellandi, Abdellah Chehri, Yurong Qian, and Gwanggil Jeon  
Volume 2016, Article ID 8312032, 3 pages

## **Zynq-Based Reconfigurable System for Real-Time Edge Detection of Noisy Video Sequences**

Iljung Yoon, Heewon Joung, and Jooheung Lee  
Volume 2016, Article ID 2654059, 9 pages

## **A Novel Seam Finding Method Using Downscaling and Cost for Image Stitching**

Jinwook Jeong and Kyungkoo Jun  
Volume 2016, Article ID 5258473, 8 pages

## **A Novel Encryption Algorithm Based on DWT and Multichaos Mapping**

Wei Wang, Haiyan Tan, Yu Pang, Zhangyong Li, Peng Ran, and Jun Wu  
Volume 2016, Article ID 2646205, 7 pages

## **Context-Aware Mobile Sensors for Sensing Discrete Events in Smart Environment**

Awais Ahmad, M. Mazhar Rathore, Anand Paul, Won-Hwa Hong, and HyunCheol Seo  
Volume 2016, Article ID 7283831, 10 pages

## **Noise Estimation and Suppression Using Nonlinear Function with *A Priori* Speech Absence Probability in Speech Enhancement**

Soojeong Lee and Gangseong Lee  
Volume 2016, Article ID 5352437, 7 pages

## **QoS and QoE Aware N-Screen Multicast Service**

Ghulam Sarwar, Farman Ullah, and Sungchang Lee  
Volume 2016, Article ID 8040138, 11 pages

## **Energy Optimization for Outdoor Activity Recognition**

Mehdi Boukhechba, Abdenour Bouzouane, Bruno Bouchard, Charles Gouin-Vallerand, and Sylvain Giroux  
Volume 2016, Article ID 6156914, 15 pages

## **Characteristics of Relocated Quiet Zones Using Virtual Microphone Algorithm in an Active Headrest System**

Seokhoon Ryu and Young-Sup Lee  
Volume 2016, Article ID 5185242, 9 pages

## **An Efficient Image Enlargement Method for Image Sensors of Mobile in Embedded Systems**

Hua Hua, Xiaomin Yang, Binyu Yan, Kai Zhou, and Wei Lu  
Volume 2016, Article ID 6370845, 12 pages

## **CRAM: A Conditioned Reflex Action Inspired Adaptive Model for Context Addition in Wireless Sensor Networks**

Majid Hussain, Muhammad Farrukh Shafeeq, Sana Jabbar, Ali Hammad Akbar, and Shahzad Khalid  
Volume 2016, Article ID 6319830, 24 pages

**Development and Coverage Evaluation of ZigBee-Based Wireless Network Applications**

Fei Ding and Aiguo Song

Volume 2016, Article ID 2943974, 9 pages

**Length Variation Effect of the Impulse Response Model of a Secondary Path in Embedded Control**

Young-Sup Lee, Yunseon Choi, and Jeakwan Kim

Volume 2016, Article ID 8270121, 7 pages

**A Smart Gateway Architecture for Improving Efficiency of Home Network Applications**

Fei Ding, Aiguo Song, En Tong, and Jianqing Li

Volume 2016, Article ID 2197237, 10 pages

**Energy Harvesting and Information Transmission Protocol in Sensors Networks**

Xue-Fen Zhang and Chang-Chuan Yin

Volume 2016, Article ID 9364716, 5 pages

**Random Secure Comparator Selection Based Privacy-Preserving MAX/MIN Query Processing in Two-Tiered Sensor Networks**

Hua Dai, Tianyi Wei, Yue Huang, Jia Xu, and Geng Yang

Volume 2016, Article ID 6301404, 13 pages

## Editorial

# Embedded Systems for Mobile Sensors

**Marco Anisetti,<sup>1</sup> Valerio Bellandi,<sup>1</sup> Abdellah Chehri,<sup>2</sup> Yurong Qian,<sup>3</sup> and Gwanggil Jeon<sup>4</sup>**

<sup>1</sup>*Dipartimento di Informatica, Università degli Studi di Milano, Via Bramante 65, 26013 Crema, Italy*

<sup>2</sup>*School of Electrical Engineering and Computer Science, University of Ottawa, 800 King Edward Avenue, Ottawa, ON, Canada K1N 6N5*

<sup>3</sup>*Software College, Xinjiang University, 14 Sheng Li Road, Urumqi, Xinjiang 830008, China*

<sup>4</sup>*Department of Embedded Systems Engineering, Incheon National University, 119 Academy-ro, Yeonsu-gu, Incheon 406-772, Republic of Korea*

Correspondence should be addressed to Gwanggil Jeon; [gjeon@inu.ac.kr](mailto:gjeon@inu.ac.kr)

Received 21 July 2016; Accepted 22 July 2016

Copyright © 2016 Marco Anisetti et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the era of IoT, embedded systems based on mobile sensor ideas turn into increasingly compound and need understanding in various training tools such as signal processing, artificial intelligence, and multimedia communication. After a long period of computational capacity centralization, we are facing a period where the computation capacity is migrating back to the periphery of the systems and then to smart objects and sensors. Currently IoT is constituted of computationally heterogeneous devices for which the ability of executing preelaborations or complete processing is fundamental as well as the ability of doing them, keeping low power consumptions and maintaining a certain level of security. The combination of signal processing techniques for embedded systems is a principal and timely research in mobile sensors and communications. The articles contained in the present issue include both reviews and basic scientific studies focused on mobile multimedia communication using fundamental or applied signal processing methods for embedded systems and applications of IoT. This issue comprises the description of streams processing (e.g., images, video, and audio) for embedded systems involved in IoT as mobile sensors/aggregators including preprocessing for signal quality enhancement (e.g., images, noise reduction), security concerns (e.g., privacy), and evolution of communication protocols and mobile sensor network architectures.

Many challenges exist in constructing smart building/home system including interconnectivity issues between

different technologies and the need of continuous connectivity between heterogeneous sensors. A smart home gateway taking responsibility for the connection between the network layer and the ubiquitous sensor network layer plays an important role in automated home Internet of Things system. The contribution of F. Ding et al. “A Smart Gateway Architecture for Improving Efficiency of Home Network Applications” proposes an integrated access gateway providing significant flexibility for user to configure and deploy supporting multi-ubiquitous sensor network. One of the basic issues for information collection and data processing in wireless sensor networks is the network coverage. The contribution of F. Ding and A. Song “Development and Coverage Evaluation of ZigBee-Based Wireless Network Applications” presented coverage evaluation and efficient development of a ZigBee-based home network application based on adaptive weighted fusion (AWF) of all data of sensing nodes and a gateway-based architecture to reexecute packet processing and then reported to the monitoring center.

Quality of Service (QoS) parameters such as bit rate and packet losses are good indicators for optimizing network services but not for characterizing user perception called Quality of Experience (QoE). The contribution of G. Sarwar et al. “QoS and QoE Aware N-Screen Multicast Service” proposed QoS and QoE-aware adaptive mapping of N-Screen devices to multicast groups at the application layer to ensure the visual quality requirements in varying network conditions.

Power consumption is another crucial aspect in IoT and mobile sensor network systems. The contribution of A. Ahmad et al. “Context-Aware Mobile Sensors for Sensing Discrete Events in Smart Environment” proposed a smart home architecture introducing the notion of context-aware low power mobile sensors sharing the same communication model based on IoT paradigm. The author tested the performances on Hadoop and the energy consumption of the sensors showing the advantages of their solution.

The mobile phone is becoming a powerful environmental sensing unit that represents a strong support for several applications domains, ranging from traffic management to advertisement and social studies. However, the limited battery capacity of mobile devices represents a big obstacle. The contribution of M. Boukhechba et al. “Energy Optimization for Outdoor Activity Recognition” presented an approach for the recognition of users’ outdoor activities able to keep mobile resources consumption under control behaving variably in function of users’ behaviors and the remaining battery level.

Dedicated sensors can harvest the energy from the environment, but they cannot charge and discharge at the same time. The contribution of X.-F. Zhang and C.-C. Yin “Energy Harvesting and Information Transmission Protocol in Sensors Networks” studied a wireless system under energy harvesting conditions proposing a method to optimize the system outage performance via finding the optimal save ratio.

Preprocessing of sensors data is a fundamental step for improving the quality of the signal to be effectively used in the following steps of the target application (e.g., image recognition, speech recognition, and human computer interaction systems). More specifically preprocessing for noise reduction has a paramount importance. Noise in video streams impact negatively any type of processing activities aimed at detecting objects. Often object detection uses edge detection as one of the basic features for the recognition. Considering the high computational complexity of applying edge detection and filtering to 1080p video, hardware implementation on reconfigurable hardware fabric is necessary. The contribution of I. Yoon et al. “Zynq-Based Reconfigurable System for Real-Time Edge Detection of Noisy Video Sequences” proposed an embedded system utilizes dynamic reconfiguration features of Zynq SoC to perform partial reconfiguration of different filter bitstreams during run-time according to the detected noise density level in the incoming video frames.

Noise estimation algorithms are essential components of many modern mobile communications. The contribution of S. Lee and G. Lee “Noise Estimation and Suppression Using Nonlinear Function with *A Priori* Speech Absence Probability in Speech Enhancement” proposed a noise-based compensation of minimum statistics method for speech enhancement in highly nonstationary noisy environments using a nonlinear function and a priori speech absence probability. Active noise cancellation is a technique based on adaptive feed-forward control mostly to superpose an artificial sound to an unwanted disturbance noise. The contribution of Y.-S. Lee et al. “Length Variation Effect of the Impulse Response Model of a Secondary Path in Embedded Control” investigated on the length variation effect of the impulse response function

for the secondary path model in active noise control using an embedded control board.

The contribution of S. Ryu and Y.-S. Lee “Characteristics of Relocated Quiet Zones Using Virtual Microphone Algorithm in an Active Headrest System” proposed a theoretical and experimental investigation on the characteristics of the relocated zone of quiet as effective antinoise solution using virtual microphone based FxLMS algorithm suitable to be embedded in a real-time digital controller for an active headrest system.

The principal and most challenging requirement for image enlargement methods in embedded systems is keeping a good balance between performance, low computational cost, and low memory usage. The contribution of H. Hua et al. “An efficient Image Enlargement Method for Image Sensors of Mobile in Embedded Systems” proposes an efficient image enlargement method based on different kinds of features for different morphologies with different approaches and the relative learned dictionaries for efficient image representation. The authors adopt clustering approach and projection matrix on dictionaries for improving speed and memory usage.

Seaming finding is an important step for creating panorama images for smoothing the differences observed at boundaries between stitched images. The contribution of J. Jeong and K. Jun “A Novel Seam Finding Method Using Downscaling and Cost for Image Stitching” proposed an improved seam finding method based on a cost function to determine how each pixel affects stitching and a downscaled version of overlapped area to approximate a seam and then interpolate the seam to the original region.

Privacy-preserving elaboration and transmission of sensing data have drawn much attention recently. Encryption of a digital image is fundamental in applications of body area networks where the captured image may include a number of privacy issues. Some of the disadvantages of the past encryption methods are the small key space and low ability of resistance to attacks. The contribution of W. Wang et al. “A Novel Encryption Algorithm Based on DWT and Multichaos Mapping” proposes a two-step discrete wavelet transform-based encryption algorithm enhanced with multichaos matrix. The authors proved that the proposed approach shows large key space, high key sensitivity, and ability of resistance to attacks.

In addition to encryption of data while in transit, privacy-preserving data queries are crucial in wireless sensor network. The contribution of H. Dai et al. “Random Secure Comparator Selection Based Privacy-Preserving MAX/MIN Query Processing in Two-Tiered Sensor Networks” proposed a privacy-preserving MAX/MIN query processing approach based on random secure comparator selection in two-tiered sensor network denoted as RSCS-PMQ.

Context awareness is a fundamental property of a system expressing the capacity of being aware of its physical environment or situation and responding in a proactive and intelligent manner. It is fundamental in the context of wireless sensor network for opening to advanced applications involving autonomous reasoning and decision-making. The contribution of M. Hussain et al. “CRAM: A Conditioned Reflex Action inspired Adaptive Model for Context Addition

in Wireless Sensor Networks” proposed a context added system in which the actuations once performed by the system help the system itself to internally evolve enriching context repository through retrospective contexts and improving introspection.

We think that this special issue would shed light on major achievements in the area of embedded systems for mobile sensors and IoT and attract attention by the scientific community to pursue further investigations leading to the rapid development of these crucial technologies.

## **Acknowledgments**

We would like to express our appreciation to all the authors for their informative contributions and the reviewers for their support and constructive critiques in making this special issue possible.

*Marco Anisetti  
Valerio Bellandi  
Abdellah Chehri  
Yurong Qian  
Gwanggil Jeon*

## Research Article

# Zynq-Based Reconfigurable System for Real-Time Edge Detection of Noisy Video Sequences

**Iljung Yoon, Heewon Joung, and Jooheung Lee**

*Department of Electronics and Computer Engineering, Hongik University, Seoul 04066, Republic of Korea*

Correspondence should be addressed to Jooheung Lee; [joolee@hongik.ac.kr](mailto:joolee@hongik.ac.kr)

Received 25 December 2015; Accepted 10 July 2016

Academic Editor: Valerio Bellandi

Copyright © 2016 Iljung Yoon et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We implement Zynq-based self-reconfigurable system to perform real-time edge detection of 1080p video sequences. While object edge detection is a fundamental tool in computer vision, noises in the video frames negatively affect edge detection results significantly. Moreover, due to the high computational complexity of 1080p video filtering operations, hardware implementation on reconfigurable hardware fabric is necessary. Here, the proposed embedded system utilizes dynamic reconfiguration capability of Zynq SoC so that partial reconfiguration of different filter bitstreams is performed during run-time according to the detected noise density level in the incoming video frames. Pratt's Figure of Merit (PFOM) to evaluate the accuracy of edge detection is analyzed for various noise density levels, and we demonstrate that adaptive run-time reconfiguration of the proposed filter bitstreams significantly increases the accuracy of edge detection results while efficiently providing computing power to support real-time processing of 1080p video frames. Performance results on configuration time, CPU usage, and hardware resource utilization are also compared.

## 1. Introduction

Heterogeneous embedded systems have proliferated in the Internet of Things era, and stream-based applications for multimedia services are widely used in various types of portable devices. These applications are data-intensive and need high computing capability to meet the required throughput and stringent real-time constraints. On the other hand, low power consumption is also as important as throughput for these portable embedded systems since they operate often in the energy constrained environments [1].

Recently, Zynq SoC (System on Chip) platform including ARM dual-core Cortex-A9 processor with FPGA fabric has been used in the embedded systems (e.g., Advanced Driver Assistance System) to implement computationally complex signal processing algorithms by utilizing both of SW flexibility of ARM processor and parallel processing capability of reconfigurable hardware fabric [2]. For sensor processing and tracking, object classification, and assessment, edge detection is a fundamental tool in computer vision applications. However, 1080p-resolution video processing for

object edge extraction in real-time is highly time consuming and becomes computational bottleneck for ARM processors. Therefore, algorithm migration onto the FPGA fabric has been preferred to meet the performance requirements. The previous studies [3, 4] had designed drawback such as inflexibility of the implemented hardware architecture and lack of adaptation capability for time-varying incoming video signals while partial reconfiguration technique on FPGA can be highly desirable to solve these problems. In particular, as shown in the previous studies [5], corrupted images with unwanted salt-and-pepper noises caused by faulty memory cells, sensing error in the analog-to-digital conversion, or bit error in transmission degrade the performance of edge detection filters significantly.

In this paper, we implement the adaptive partial reconfigurable system to maximize the output performance of the implemented edge detection filter. By detection of noise density levels in the incoming video, adaptive self-reconfiguration of hardware bitstream is performed in real-time in order to remove unwanted noises before edge extraction process. For efficient utilization of hardware resources,



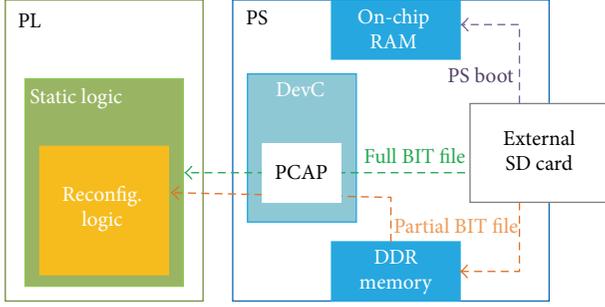


FIGURE 2: PR interface using PCAP.

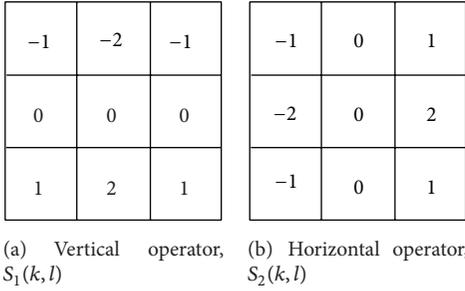


FIGURE 3: Sobel operator.

download throughput. PR interface using PCAP is shown in Figure 2. First Stage Boot Loader (FSBL) read from external SD card boots up PS and configures the PL with the full bitstream via the PCAP, and user application loads the partial bitstream into DDR memory later on. From this moment, software-controlled partial reconfiguration is enabled to dynamically reconfigure part of the PL with the bitstream of preimplemented IP core [17].

### 3. Proposed Approach

For real-time object edge extraction of 1080p-resolution video frames, Sobel filter has been implemented on PL region. As an orthogonal gradient operator, Sobel operators shown in Figure 3 are used to perform 2-dimensional convolution in every pixel of the incoming video frame.  $f(x, y)$  is the pixel value at location  $(x, y)$  [18]

$$G_1(x, y) = \sum_{k=-1}^1 \sum_{l=-1}^1 S_1(k, l) f(x+k, y+l), \quad (1)$$

$$G_2(x, y) = \sum_{k=-1}^1 \sum_{l=-1}^1 S_2(k, l) f(x+k, y+l).$$

The gradient vector magnitude and direction are given by

$$G(x, y) = \sqrt{G_1^2(x, y) + G_2^2(x, y)}, \quad (2)$$

$$\theta_G(x, y) = \tan^{-1} \frac{G_2(x, y)}{G_1(x, y)}.$$

Typically, salt-and-pepper noise is caused by defective sensors, faulty memory cells, and bit error in transmission, and it degrades the performance of edge detection filter significantly. In this paper, we implement the noise detection algorithm proposed in [19] and briefly describe it as follows.

$p(x, y)$  is a pixel value to be processed, and  $3 \times 3$  window with  $p(x, y)$  as the center location is considered.  $p_{\min}$  and  $p_{\max}$  are minimum and maximum pixel values of  $3 \times 3$  window. Then, thresholds  $T_{\min}$  and  $T_{\max}$  are defined as

$$T_{\min} = \begin{cases} 2p_{\min} - 255, & 2p_{\min} > 255, \\ 0, & \text{otherwise,} \end{cases} \quad (3)$$

$$T_{\max} = \begin{cases} 2p_{\max}, & 2p_{\max} < 255, \\ 255, & \text{otherwise.} \end{cases}$$

Equation (4) is used as a criterion to determine whether  $p(x, y)$  is a corrupted noise pixel or not

$$\text{noise}_{\text{det}} = \begin{cases} 1, & p(x, y) \leq T_{\min} \text{ or } p(x, y) \geq T_{\max}, \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

Then, the noise density is measured as the total number of detected noise pixels divided by the total number of pixels in a given video frame.

Since the edge detection performance decreases significantly in the corrupted image by the salt-and-pepper noise, Median filter is implemented for denoising as in

$$\widehat{M}(x, y) = \text{Median} \{M_{x', y'} : (x', y') \in W(x, y)\}. \quad (5)$$

Here, Median value of neighboring pixels in window  $W$  is selected as output [20].

The proposed self-reconfiguration method replaces PRMs to the Sobel edge detection after preprocessing Median operator (hereafter referred to as the Median + Sobel filter) which is effective for noise reduction when the salt-and-pepper noise is added to the video frame.

Pratt's Figure of Merit (FOM) to evaluate the accuracy of detected edge in noisy image is used to determine corresponding threshold of noise density level. As performance of edge detection accuracy is deteriorated, partial reconfiguration process is triggered to reduce noise before edge filtering. Pratt's FOM is defined by

$$R = \frac{1}{I_N} \sum_{i=1}^{I_D} \frac{1}{1 + \alpha d^2}. \quad (6)$$

Here,  $I_N = \max(I_I, I_D)$ ,  $I_I$  is the number of edge points in the ideal edge,  $I_D$  is the number of edge points in the detected edge,  $\alpha$  is a calibration constant, and  $d$  is the distance between the detected and the ideal edges [21]. The distance " $d$ " is important factor in the evaluation of edge detection using PFOM. The factor " $d$ " is inversely proportional to the factor  $R$ . For a stained edge, the distance " $d$ " between ideal and detected edge increases and factor  $R$  is reduced.

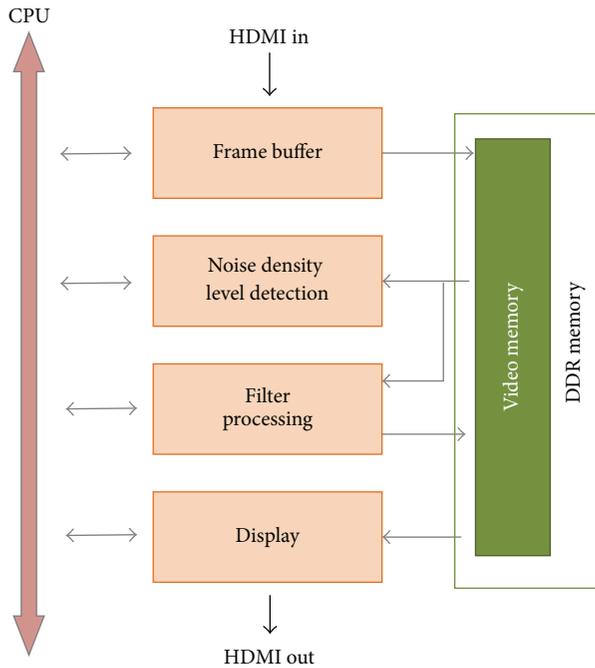


FIGURE 4: Video pipeline and noise detection task.

Figure 4 shows video pipeline architecture and noise detection task. The 1080p video frames from HDMI-IN are stored in DDR memory. The implemented filter mainly consists of three subfunctions that are filtering process, edge detection process, and bus interface to control input and output of video. Video DMA (VDMA) reads video frames from DDR memory and sends them to the filter engine. The AXI4-Stream interfaces are connected through VDMA and AXI interconnect block to the high-performance slave port of the PS [22]. The output frame from the filter engine is stored back into DDR memory and then stored frame is sent to the display controller for HDMI-OUT. Synthesized results show that pipelined edge detection process has 9 clock cycles of latency, and total processing time requires 2,059 clock cycles.

Noise density level detection is performed to trigger partial reconfiguration of Median + Sobel filter. Partial bitstreams for filter operations are loaded from SD card into DDR memory by the user application running on the PS. It improves the reconfiguration time and also takes advantage of caching. Then, the application can use partial bitstreams to modify the partially reconfigurable region in PL without interrupting the rest of the PL area. Partial reconfiguration of Sobel or Median + Sobel bitstreams from DDR memory to the predefined PL region is performed through the PCAP interface. If the measured noise density becomes higher than the threshold, then Median + Sobel bitstream is configured to the partially reconfigurable region (PRR) to replace Sobel bitstream.

For Zynq SoC, an AXI-PCAP bridge consisting of “transmit” and “receive” FIFO buffers between the AXI and the PCAP interface is implemented in the Device Configuration interface (DevC) of the PS. This bridge converts 32-bit AXI formatted data to 32-bit PCAP protocol, and a DMA engine

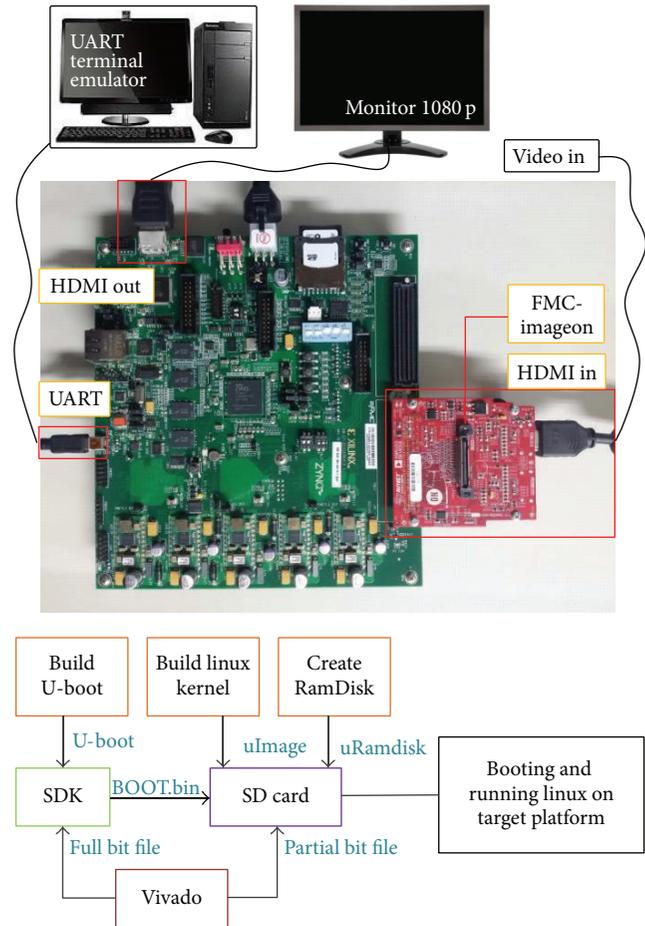


FIGURE 5: Experimental environment.

transfers data between the FIFOs and the DDR memory for partial reconfiguration. A DevC driver function, built on top of sysfs, is called to move data across the PCAP interface through initiating the DMA transaction and then waits for an interrupt signaling that the transfer is completed. When both AXI and PCAP transfers are finished, then the function call returns. The application does not need to know about physical location of partially reconfigurable region because partial bitstream has the configuration frame addressing information. The filter in PL region is reset before transferring a partial bitstream via DevC/PCAP. When the bitstream transfer is completed, the reset is released and the configured filter is restarted with VDMA. Our measurements show that up to 5 frames of incoming video can be dropped during the partial reconfiguration process.

## 4. Experimental Results

Devices used in the experiment are ZC702 evaluation board with XC7Z020 AP SoC, FMC module equipped with HDMI input/output based on ADV7611/ADV7511, and 1920 × 1080 resolution monitor. ZC702 board is controlled through UART Terminal Emulator running on PC [23]. Figure 5 shows experimental setup for implementation of the proposed

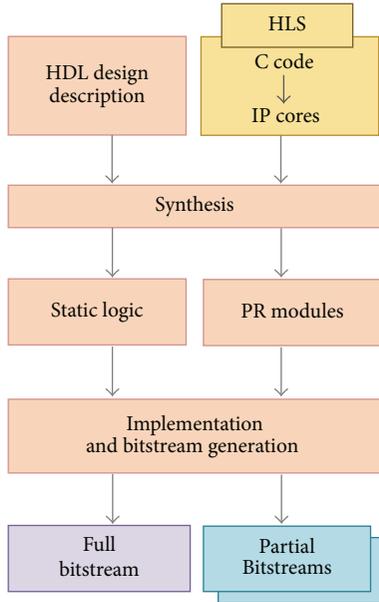


FIGURE 6: The procedure of bitstream generation.

reconfigurable edge detection system. The Boot binary file booting the ZynqSoC consists of Zynq FSBL created in the SDK tool, full bit file generated in the Vivado, and U-boot called second stage boot loader. The compressed kernel image, that is, ulmage, supports linux operating system on the target board [24]. The partial bit files are initially stored in SD card and read to DDR memory for PS to perform PR through the PCAP interface. The target board utilizes uramdisk as the root file system. The Software Development Kit (SDK) tool is used to create linux application on the processor to perform the operation of the proposed method and partial reconfiguration.

In this paper, the proposed Sobel and Median + Sobel filter blocks are implemented by High-Level Synthesis (HLS) Tool [25, 26]. The HLS tool transforms C language, C++, and SystemC into a RTL implementation, and also offers the pipelining of the function through GUI interface.

Vivado Integrated Design Environment (IDE) is a development tool to provide Xilinx Integrated Synthesis Environment (ISE) and Xilinx Platform Studio (XPS). It is used to analyze and synthesize the HDL designs and perform timing analysis. Figure 6 shows the overall procedure of full and partial bitstream generation. The HDL design description of the system and the IP cores generated by the HLS tool are synthesized. Then, as shown in Figure 7, we floorplanned partially reconfigurable region so that hardware resources required for implementation of PRMs are less than 90% of the total amount of the PRR hardware resources. The hardware resource comparison of PRR and PRMs is summarized in Table 1.

As a result, one full bitstream and two partial bitstreams are generated. The PL system is initially configured with a full bitstream including static logic and Sobel filter. If detected noise density level becomes higher than threshold, partial bitstream of Median + Sobel is used to reconfigure the PRR

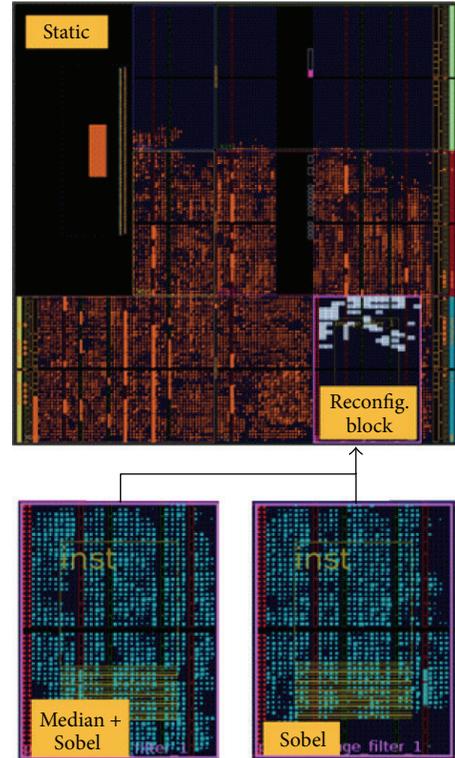


FIGURE 7: Static logic and PR modules.

TABLE 1: Comparison of PRR and PRMs resources.

Resources	PRR	PRMs	
	Available	Sobel	M + S
LUT	7800	2871 (36.8%)	3478 (44.6%)
SLICE	1400	1054 (75.3%)	1220 (87.1%)
RAMB18	40	3 (7.5%)	6 (15.0%)
DSPs	40	0 (0%)	0 (0%)

region. If its level becomes lower than threshold, partial bitstream of Sobel is read again from DDR memory for runtime reconfiguration.

In this paper, new reconfiguration interface called PCAP (Processor Configuration Access Port) available on Xilinx Zynq SoC is explored to perform partial reconfiguration of filter bitstreams using ARM Cortex-A9 processor. While theoretical speed of reconfiguration for 32-bit PCAP interface clocked at 100 MHz is 400 MB/s, practical reconfiguration speed is much lower due to the internal ARM interconnect architecture. Several examples using filter bitstream with JTAG, ICAP, and PCAP interfaces are shown in Table 2.

Due to the design approach using partial reconfiguration, we could achieve significant reduction of both bitstream file size and reconfiguration time through PCAP interface. As shown in Table 3, partial reconfiguration time is reduced to 12% of the full configuration time, and system downtime to replace the function of the proposed filter engine is not necessary any longer.

TABLE 2: Different configuration interfaces.

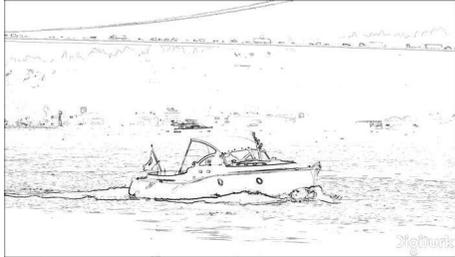
	[27]	[28]	[29]	[30]	Proposed
Partial bitstream size	242 KB	244 KB	100 KB	374 KB	460 KB
Partial reconfig. time	1 sec	460 ms	29 ms	6 ms	10 ms
Configuration interface	JTAG (Virtex-4)	ICAP (Virtex-4)	ICAP (Virtex-5)	PCAP (Zynq-7000)	PCAP (Zynq-7000)
Frequency (MHz)	6 MHz	100 MHz	100 MHz	100 MHz	100 MHz
Types of module	Median filter	Median filter	Electronic stethoscope	FPU	Median + Sobel filter



(a) Noisy video #1 (salt-and-pepper noise density: 20%)



(b) Noisy video #2 (salt-and-pepper noise density: 20%)



(c) Output of Sobel filter: original video #1



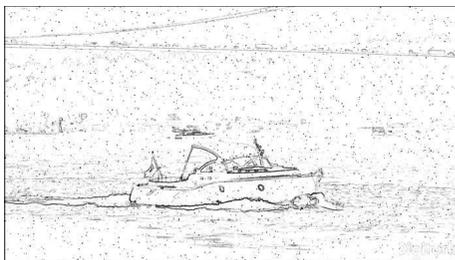
(d) Output of Sobel filter: original video #2



(e) Output of Sobel filter (noisy video #1)



(f) Output of Sobel filter (noisy video #2)



(g) Output of Median + Sobel filter (noisy video #1)



(h) Output of Median + Sobel filter (noisy video #2)

FIGURE 8: Comparison of edge detection results for the noisy video sequences.

Full HD video sequences of  $1920 \times 1080$  resolution are used for experiments. Boat sequence (video #1) and Beauty sequence (video #2) are shown in Figure 8. Figures 8(a) and 8(b) show video sequences #1 and #2 corrupted by salt-and-pepper noise with 20% noise density level, respectively. Sobel

filter's outputs of original video sequences #1 and #2 are shown in Figures 8(c) and 8(d). Edges in the scene are clearly detected. In contrast, Figures 8(e) and 8(f) show that Sobel filter's performance for the noisy video sequences #1 and #2 degrades significantly, resulting in lower values of PFOMs.

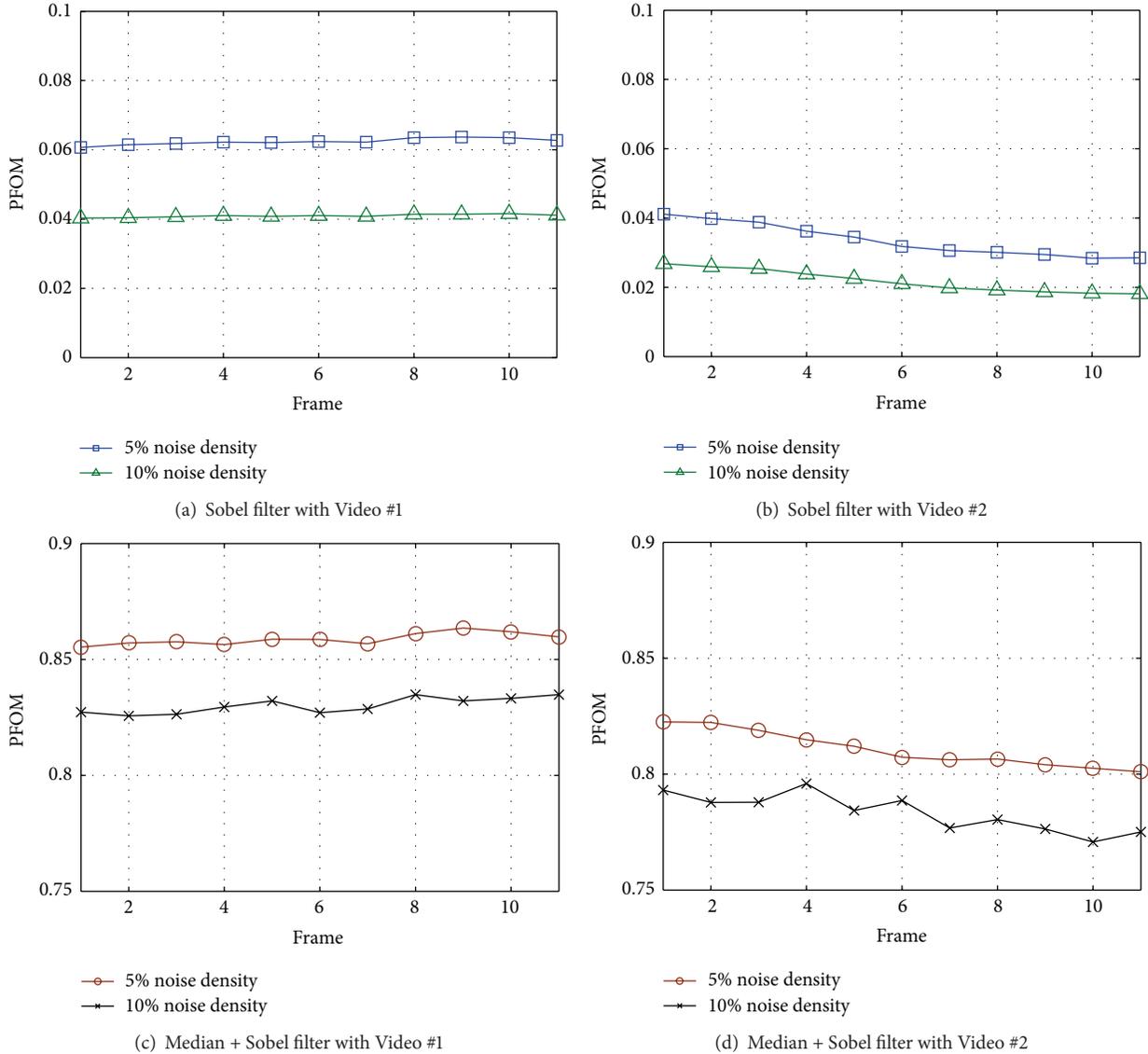


FIGURE 9: PFOMs for Sobel and Median + Sobel filters.

TABLE 3: Bitstream size and configuration time.

	Full bitstream	Partial bitstream
Bitstream size	4,045,564 bytes	460,544 bytes
PCAP configuration time	83 ms	10 ms

As shown in Figures 8(g) and 8(h), Median + Sobel filter is highly effective for the noisy video sequences #1 and #2. Its edge detection results are significantly improved both subjectively and objectively. For objective evaluation of edge detection results, PFOM is used to compare the performance of Sobel and Median + Sobel filters [31, 32]. Video sequences #1 and #2 are corrupted by salt-and-pepper noise with 5% and 10% noise density level. Since Median operator removes the salt-and-pepper noise in the corrupted video frames, performance analysis of two filter engines shows that Median + Sobel filter provides about 14 to 20 times improvement of PFOM as indicated in Figure 9.

In Figure 10, frame rates supported by Sobel and Median + Sobel filters are indicated, and run-time CPU usage of hardware and software filter implementations is measured in Figure 11. While 100% of CPU power is used for S/W implementation of Sobel filter, frame rates drop significantly down to 1.5 fps, indicating software implementation is not suitable for real-time processing of 1080p video frames. Here, camera controller supports 60 input frames per second. Due to the additional computational complexity, H/W Median + Sobel filter supports up to 29 frames per second (fps), about 1 frame less than Sobel H/W implementation (30 frames per second).

After PR, the power consumption of hardware platform on Xilinx Zynq FPGA is estimated using Power Report in Vivado Design Suite [33].

As shown in Figure 12, the power consumption of Median + Sobel filter is higher than Sobel filter because Median + Sobel filter requires more hardware resources in FPGA than Sobel filter.

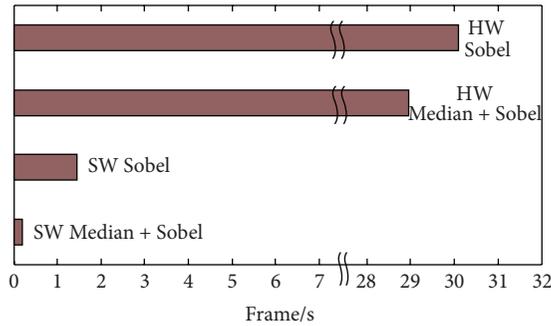


FIGURE 10: Comparison of frame rates.

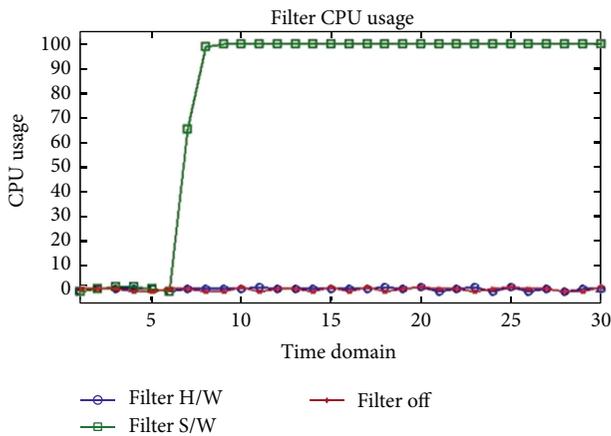


FIGURE 11: CPU Usage.

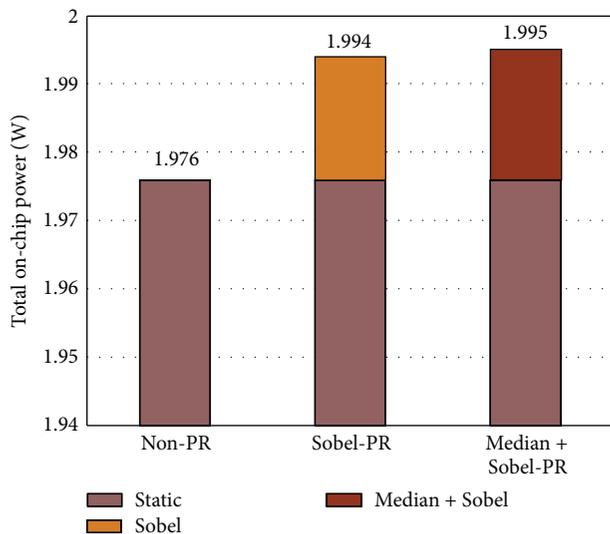


FIGURE 12: Comparison of power consumption.

## 5. Conclusion

In this paper, we propose adaptive partial reconfigurable system to maximize the output performance of the implemented edge detection filter. Hardware implementation of filter engine onto the FPGA fabric provides computing capability of real-time edge detection of 1080p video sequences.

According to detection of noise density levels in the incoming video, adaptive self-reconfiguration of hardware bitstream is performed during run-time and it enables significant performance improvement in both subjective and objective results. Experimental results show that partial reconfiguration time is reduced to 12% of the full configuration time, and about 14 to 20 times improvement of PFOM is achieved.

## Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (NRF-2012R1A1A1008674) and also by Business for Cooperative R&D between Industry, Academy, and Research Institute funded Korea Small and Medium Business Administration (C03319350100437169) and 2014 Hongik University Research Fund. Additionally, the authors would like to acknowledge Xilinx and the Xilinx University Program for its generous donation of S/W design tools and H/W boards.

## References

- [1] A. Majumdar, S. Cadambi, and S. T. Chakradhar, "An energy-efficient heterogeneous system for embedded learning and classification," *IEEE Embedded Systems Letters*, vol. 3, no. 1, pp. 42–45, 2011.
- [2] UG1165 (v2015.3), *Zynq-7000 All Programmable SoC: Embedded Design Tutorial*, Xilinx, November 2015.
- [3] D. Crookes, K. Benkrid, A. Bouridane, K. Aiotaibi, and A. Benkrid, "Design and implementation of a high level programming environment for FPGA-based image processing," *IEE Proceedings: Vision, Image and Signal Processing*, vol. 147, no. 4, pp. 377–384, 2000.
- [4] P. Greisen, M. Runo, P. Guillet et al., "Evaluation and FPGA implementation of sparse linear solvers for video processing applications," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 8, pp. 1402–1407, 2013.
- [5] A. M. Mahmood, H. H. Maraş, and E. Elbaşı, "Measurement of edge detection algorithms in clean and noisy environment," in *Proceedings of the 8th IEEE International Conference on Application of Information and Communication Technologies (AICT '14)*, pp. 1–6, Astana, Kazakhstan, October 2014.
- [6] Xilinx, *Zynq-7000 All Programmable SoC Overview*, DS190(v1.8), Xilinx, 2015.
- [7] UG585(v1.10), *Zynq-7000 All Programmable SoC Technical Reference Manual*, Xilinx, February 2015.
- [8] J. Silva, V. Sklyarov, and I. Skliarova, "Comparison of on-chip communications in Zynq-7000 all programmable systems-on-chip," *IEEE Embedded Systems Letters*, vol. 7, no. 1, pp. 31–34, 2015.
- [9] Xilinx, *Vivado Design Suite User Guide Partial Reconfiguration*, UG909(v2014.4), Xilinx, 2014.
- [10] E. Stott, P. Sedcole, and P. Y. K. Cheung, "Fault tolerant methods for reliability in FPGAs," in *Proceedings of the International*

- Conference on Field Programmable Logic and Applications (FPL '08)*, pp. 415–420, Heidelberg, Germany, September 2008.
- [11] C. Insaurralde, “Reconfigurable computer architectures for dynamically adaptable avionics systems,” *IEEE Aerospace and Electronic Systems Magazine*, vol. 30, no. 9, pp. 46–53, 2015.
- [12] M. G. Parris, C. A. Sharma, and R. F. Demara, “Progress in autonomous fault recovery of Field Programmable Gate Arrays,” *ACM Computing Surveys*, vol. 43, no. 4, article 31, 2011.
- [13] M. Liu, W. Kuehn, Z. Lu, and A. Jantsch, “Run—time partial reconfiguration speed investigation and architectural design space exploration,” in *Proceedings of the International Conference on Field Programmable Logic and Application*, pp. 498–502, Prague, Czech Republic, September 2009.
- [14] R. Bonamy, H.-M. Pham, S. Pillement, and D. Chillet, “UPaRC—Ultra-fast power-aware reconfiguration controller,” in *Proceedings of the Design, Automation & Test in Europe Conference & Exhibition (DATE '12)*, pp. 1373–1378, IEEE, Dresden, Germany, March 2012.
- [15] M. Hübner, D. Göhringer, J. Noguera, and J. Becker, “Fast dynamic and partial reconfiguration data path with low hardware overhead on Xilinx FPGAs,” in *Proceedings of the IEEE International Symposium on Parallel & Distributed Processing*, pp. 1–8, April 2010.
- [16] K. Vipin and S. A. Fahmy, “Zycap: efficient partial reconfiguration management on the xilinx zynq,” *IEEE Embedded Systems Letters*, vol. 6, no. 3, pp. 41–44, 2014.
- [17] XAPP1159(v1.0) and C. Kohn, *Partial Reconfiguration of a Hardware Accelerator on Zynq-7000 All Programmable SoC Devices*, Xilinx, January 2013.
- [18] S. Jin, W. Kim, and J. Jeong, “Fine directional de-interlacing algorithm using modified Sobel operation,” *IEEE Transactions on Consumer Electronics*, vol. 54, no. 2, pp. 857–862, 2008.
- [19] P.-Y. Chen, C.-Y. Lien, and Y.-M. Lin, “A real-time image denoising chip,” in *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS '08)*, pp. 3390–3393, Seattle, Wash, USA, May 2008.
- [20] C. Chen, J. Ni, and J. Huang, “Blind detection of median filtering in digital images: a difference domain based approach,” *IEEE Transactions on Image Processing*, vol. 22, no. 12, pp. 4699–4710, 2013.
- [21] W. K. Pratt, *Digital Image Processing*, PIKS Inside, 3rd edition, 2000.
- [22] UG1037 (v3.0), *Vivado Design Suite : AXI Reference Guide*, Xilinx, June 2015.
- [23] C. Kohn, *Partial Reconfiguration of a Hardware Accelerator with Vivado Design Suite for Zynq-7000 ApSoC Processor*, XAPP1231 (v1.1), Xilinx, 2015.
- [24] UG821 (v12.0), *Zynq-7000 All Programmable SoC Software Developers Guide*, Xilinx, San Jose, Calif, USA, 2015.
- [25] F. M. Vallina, C. Kohn, and P. Joshi, *Zynq All Programmable SoCSobel Filter Implementation Using the Vivado HLS Tool*, XAPP890 (v1.0), Xilinx, 2012.
- [26] Xilinx, *Vivado Design Suite User Guide: High-Level Synthesis*, UG902 (v2015.4), Xilinx, 2015.
- [27] S. U. Bhandari, S. Subbaraman, S. S. Pujari, and R. Mahajan, “Real time video processing on FPGA using on the fly partial reconfiguration,” in *Proceedings of the International Conference on Signal Processing Systems (ICSPS '09)*, pp. 244–247, Singapore, May 2009.
- [28] S. U. Bhandari, S. Subbaraman, S. Pujari, and R. Mahajan, “Internal dynamic partial reconfiguration for real time signal processing on FPGA,” *Indian Journal of Science and Technology*, vol. 3, no. 4, pp. 365–368, 2010.
- [29] A. A. Prince and S. Mishra, “Multi-mode electronic stethoscope implementation and evaluation using Dynamic Reconfigurable Design,” in *Proceedings of the 5th IEEE International Advance Computing Conference (IACC '15)*, pp. 228–232, Bangalore, India, June 2015.
- [30] M. Al Kadi, P. Rudolph, D. Göhringer, and M. Hübner, “Dynamic and partial reconfiguration of Zynq 7000 under Linux,” in *Proceedings of the International Conference on Reconfigurable Computing and FPGAs (ReConFig '13)*, pp. 1–5, IEEE, Cancun, Mexico, December 2013.
- [31] I. E. Abdou and W. K. Pratt, “Quantitative design and evaluation of enhancement/thresholding edge detectors,” *Proceedings of the IEEE*, vol. 67, no. 5, pp. 753–763, 1979.
- [32] J.-A. Jiang, C.-L. Chuang, Y.-L. Lu, and C.-S. Fahn, “Mathematical-morphology-based edge detectors for detection of thin edges in low-contrast regions,” *IET Image Processing*, vol. 1, no. 3, pp. 269–277, 2007.
- [33] Xilinx, *Vivado Design Suite User Guide: Power Analysis and Optimization*, UG907(v2015.4), Xilinx, 2015.

## Research Article

# A Novel Seam Finding Method Using Downscaling and Cost for Image Stitching

**Jinwook Jeong and Kyungkoo Jun**

*Department of Embedded Systems Engineering, University of Incheon, Incheon 22012, Republic of Korea*

Correspondence should be addressed to Kyungkoo Jun; [kjun@incheon.ac.kr](mailto:kjun@incheon.ac.kr)

Received 23 January 2016; Accepted 18 April 2016

Academic Editor: Yurong Qian

Copyright © 2016 J. Jeong and K. Jun. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Seaming finding is an important step for creating panorama images because it smoothes away differences observed at boundaries between stitched images. We propose an improved seam finding method in which we define a cost function to measure the discrepancies that boundary pixels cause. We are also able to improve computing complexity by avoiding finding a seam over the whole area of overlapped region. Instead, we use a downscaled version of overlapped area to approximate a seam and then interpolate the seam to the original region. From experiments to generate panorama images, we compare our method with three other existing seam finding algorithms and observe that our method is able to produce better quality panorama image than the existing methods, while the processing time is comparable to those of the others.

## 1. Introduction

Panorama images contribute to providing realistic experience by having a wider field of view and have various application areas such as CCTV and action cams [1–3]. In general, cameras equipped with wide angle lenses are used to capture panorama images, or postprocessing software creates panorama by stitching multiple images taken by the same or different cameras.

Postprocessing software for panorama generation consists of two steps. The first step is called registration in which feature points are detected and matched between neighbor images to compute parameters such as focal distance and homography. These parameters are then used to warp images to create a large panorama. The second step is blending which balances exposure differences between images and removes seams and ghost effects from moving objects, aiming to remove stitching trace as much as possible [4–7].

The steps involved in panorama creation depend on various parameter settings, making it harder to create complete and perfect panorama images. Diverse research efforts have been made to improve the quality of resulting panorama. Various blending methods have been proposed [8–11]. A technique based on camera motion estimation was proposed

to improve the performance of the registration [8]. Seam finding algorithm was proposed by using image patches obtained by graph-cut algorithm [9]. The performance of seam finding was improved by image segmentation based on watershed algorithm [10]. Dynamic programming technique was able to improve both speed and precision even though it worked on pixel level [11].

This paper proposes a seam finding method for the blending step, which is the second step of the panorama creation. We aim to improve performance while preserving the precision of seam location. For performance, we scale down images by resizing them to half their size, resulting in less computation time. The seam line that we find in the small sized images is just the approximation of the perfect seam line. Based on the approximation, since the region where the seam line exists is known, we can limit the area where seam finding should be performed rather than applying to whole area. It results in improved performance, while maintaining precision.

This paper is organized as follows. Section 2 describes the proposed seam finding method in detail and Section 3 evaluates the performance of the method and compares it with those of existing methods. And Section 4 concludes the paper.

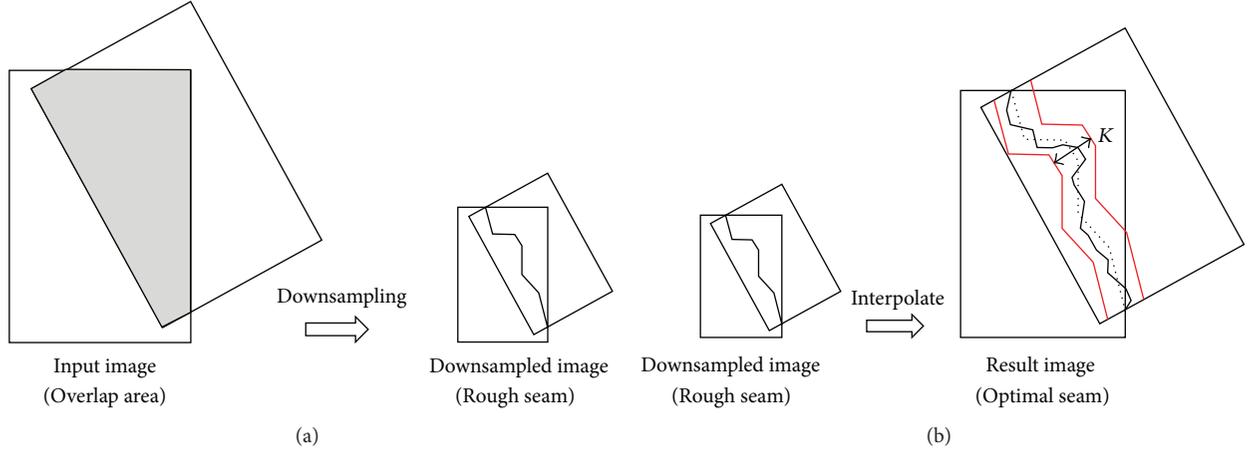


FIGURE 1: Seam finding steps using enhanced dynamic programming: (a) step 1 and (b) step 2.

## 2. Downscaling and Cost Function Based Seam Finding Method

We propose using downsampled images for estimating seam because it is able to reduce time. But it is challenging not to sacrifice precision of resulting seams. The proposed scheme consists of two sequential steps. The first step is to estimate an approximated seam from an overlapped region between two images as shown in Figure 1(a). Note that, before starting seam finding by dynamic programming, we reduce the size of the overlapped area by  $r$  times. Thus it has the effect to reduce the computation load and time required for seam finding. However, the resulting seam is just an approximation close to the real seam. In the second step shown in Figure 1(b), the reduced region is restored back to the original size. During this phase, the seam is also extended, filling up missing parts by interpolation. With the interpolated seam, we set a new area which is within  $K$  pixels away from the seam. Then, we are sure that a real seam is contained in the area. Finally, we perform the dynamic programming again over the area to find the real seam:

$$e_i = \sqrt{(G_i)^2 + (I_i)^2},$$

$$G_i = \frac{\|\nabla P_1^i + \nabla P_2^i\|}{\max(\|\nabla P_1^i + \nabla P_2^i\|)}, \quad (1)$$

$$I_i = \frac{\|P_1^i - P_2^i\|}{\max(\|P_1^i - P_2^i\|)},$$

$$E_{(i,j)} = e_{(i,j)} + \min(E_{(i-1,j-1)}, E_{(i,j-1)}, E_{(i+1,j-1)}, E_{(i-1,j)}, E_{(i+1,j)}). \quad (2)$$

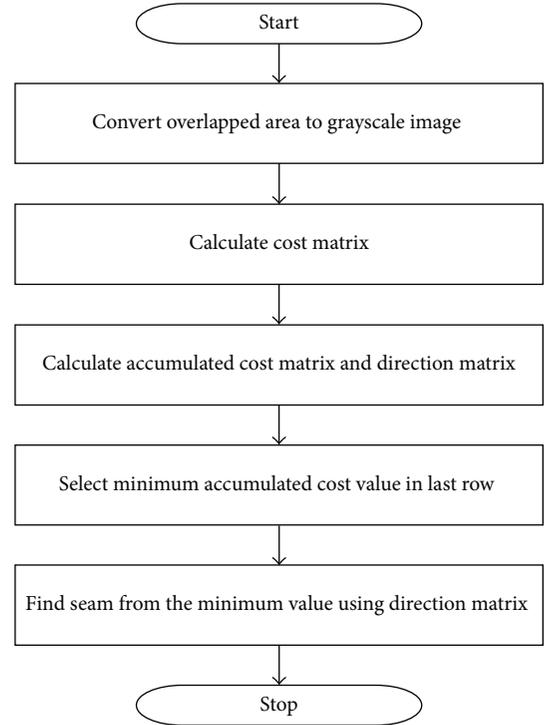


FIGURE 2: Flow chart of the seam finding.

Figure 2 is the flow chart of the seam finding. We find a seam between two images,  $Img_1$  and  $Img_2$ . At first, an overlapped region between  $Img_1$  and  $Img_2$  is determined. Then, the subimage of  $Img_1$  is located which belongs to the overlapped region and it is converted to gray scale. We call it  $Gray_1$ . In the same way, we locate  $Gray_2$ , the subimage from  $Img_2$ . It is obvious that there exists a corresponding pixel within  $Gray_2$  for every pixel from  $Gray_1$ . We call a pixel  $i$  of  $Gray_1$  and its corresponding  $Gray_2$  pixel as  $P_1^i$  and  $P_2^i$ ,

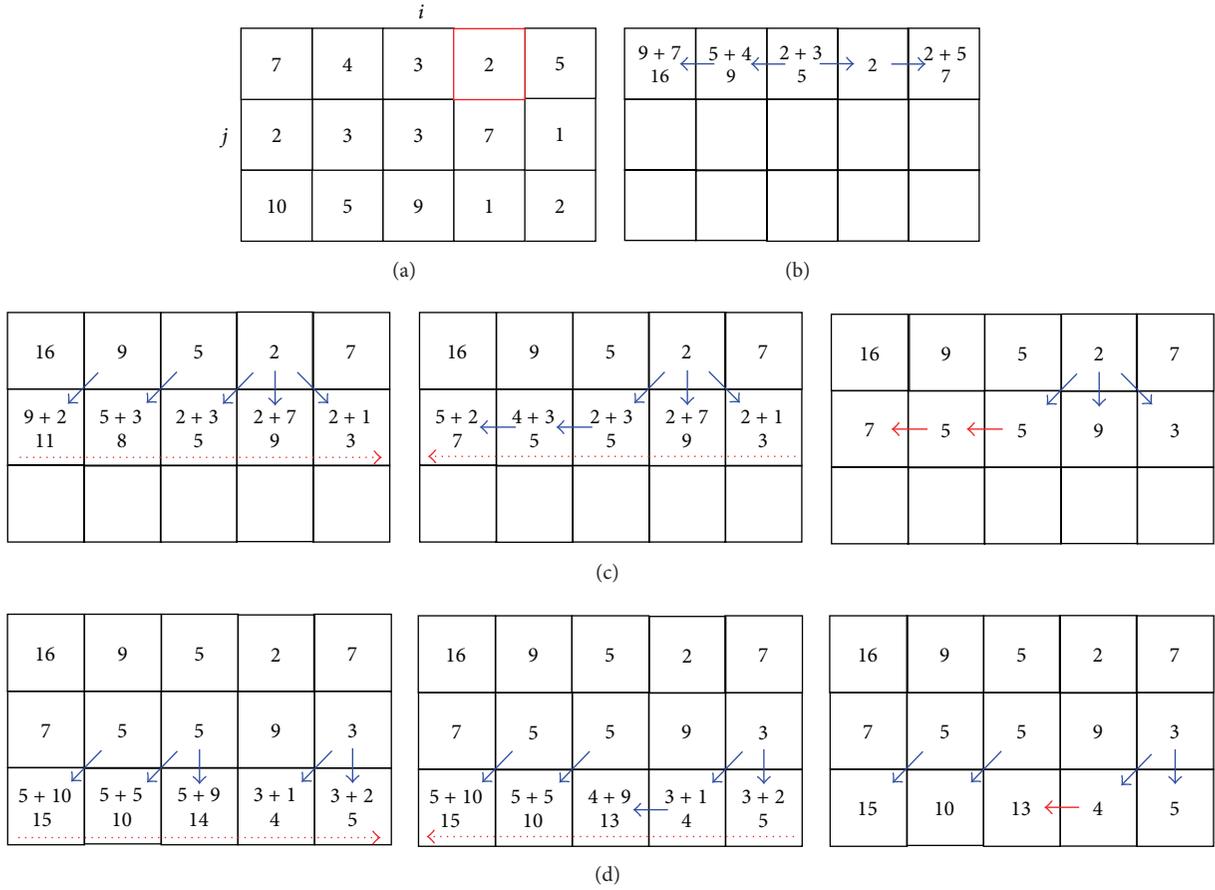


FIGURE 3: Process of accumulated cost matrix calculation (a)  $M_{cost}$ , (b) first row calculation of  $M_{a\_cost}$ , (c) second row calculation of  $M_{a\_cost}$ , and (d) third row calculation of  $M_{a\_cost}$ .

respectively. And we call  $P_i$  the pixel of the overlapped region which has corresponding pixels  $P_1^i$  and  $P_2^i$ .

Then, for every pixel  $P_i$ , we compute its cost  $e_i$ . The cost, also called energy, is a real value assigned for each pixel and represents how important the pixel is. Often, gradient or intensity of a pixel is used as the cost. In this paper, we compute  $e_i$  as (1), where  $\nabla P_1^i$  and  $\nabla P_2^i$  are the gradients of  $P_1^i$  and  $P_2^i$ .

Then, we can obtain a cost matrix  $M_{cost}$  for the overlapped region. Each element of  $M_{cost}$  corresponds to  $P_i$  and has  $e_i$  as its value. From  $M_{cost}$ , two other matrices are computed. One is an accumulated cost matrix  $M_{a\_cost}$  and the other is a direction matrix  $M_{dir}$ . By using information from both matrices, we are able to determine where a seam can be drawn without noticeable differences between  $Img_1$  and  $Img_2$ .  $M_{a\_cost}$  is used to find candidate pixels on which a seam is drawn, while  $M_{dir}$  provides pixel-by-pixel directional information of the seam.

Figure 3 shows how  $M_{a\_cost}$  matrices are made from  $M_{cost}$ . Each element of  $M_{a\_cost}$ ,  $E(i, j)$ , is called an accumulated cost of a pixel at  $i$ th column and  $j$ th row. It is calculated as (2), where  $e(i, j)$  is the cost of the pixel at  $(i, j)$  from  $M_{cost}$ .  $E(i, j)$  of a pixel is computed by adding its cost to

the minimum accumulated cost of its neighboring pixels. Note that, among eight neighboring pixels, the three at the lower row are ignored from consideration. Also, the top row pixels do not consider the upper row pixels because they do not exist.

For ease of understanding, we describe step by step how each  $E(i, j)$  is computed. At start, among the top row pixels, we find the pixel with the minimum  $e(i, j)$  and set  $E(i, j) = e(i, j)$  which is the fourth pixel with the value 2 in Figure 3(a). Then, accumulated costs of neighbor pixels  $E(i - 1, j)$  and  $E(i + 1, j)$  are computed according to (2). This procedure is repeated in a flooding way until all the pixels at the top row are processed, resulting in Figure 3(b). We then move on to the next row. Note that when calculating  $E(m, n)$ , still  $E(m + 1, n)$  is not determined yet. Thus we need to iterate twice; however the direction of calculation changes: from left to right first and then from right to left as shown in Figure 3(c): the leftmost figure of Figure 3(c) shows how the calculation proceeds from left to right, the middle one shows that the calculation progresses from right to left, and the leftmost one is the result of the calculation. Figure 3(d) shows how the third row is processed. The whole process continues until the last row is finished.

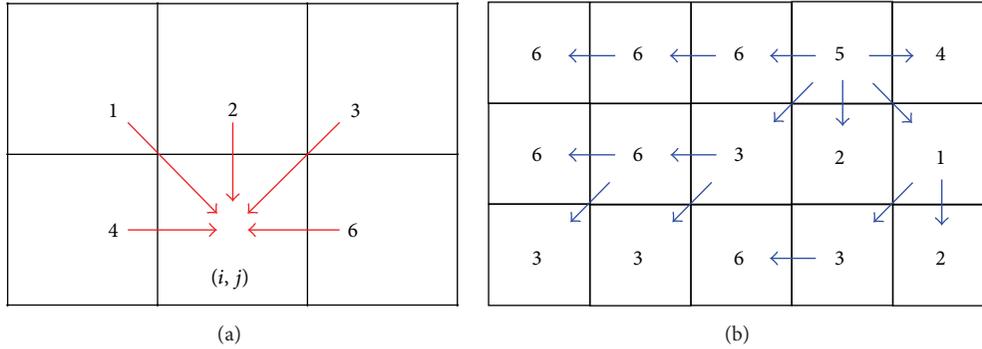


FIGURE 4: Direction description of  $M_{dir}$ : (a) direction of  $M_{dir}$  and (b) example of  $M_{dir}$ .

With  $M_{a\_cost}$ , we can determine for  $E(i, j)$  which neighboring pixel is selected to contribute to the computation. Figure 4(a) shows five candidate pixels each given numbers 1 to 6, except 5 which is reserved for the first pick pixel at the top row.  $M_{dir}$  stores for each pixel which neighboring pixel is selected for its  $E(i, j)$ . For example, Figure 4(b) shows an example  $M_{dir}$  corresponding to Figure 3(d). Note that the fourth pixel at the top row has the value 5, because no neighboring pixel is considered. The arrows are drawn for ease of understanding.

Once completing  $M_{a\_cost}$  and  $M_{dir}$ , a seam can be determined. It starts from the bottom row and moves up in an antiarrow direction, drawing a seam over visited pixels. To describe the procedure in detail, among the pixels on the last row of  $M_{a\_cost}$ , we select the one with the minimum  $E(i, j)$ . Then, we follow the arrow in a reverse way to move to one of the neighboring pixels; for this we use the direction information from  $M_{dir}$ . Finally, if we connect all the pixels that were visited before reaching the top row, those pixels consist of a seam. Since the seam consists of only the pixels with least distinction, the resulting boundary provides smooth transition between two images.

There is an unusual case in which more than one pixel has the minimum cost at the top row in  $M_{cost}$ . In this case, we cannot avoid increasing computation load, because, for each pixel, we compute  $M_{a\_cost}$  and  $M_{dir}$  independently and separately. Thus, we produce multiple sets of resulting matrices. However, we choose only the set of which  $M_{a\_cost}$  contains the least value  $E(i, j)$  at its bottom row, ignoring the rest. Then, the seam determination is processed in the same way as explained.

### 3. Performance Evaluation

We evaluate the performance of the proposed scheme through experiments that create panorama images by stitching multiple overlapping images. In the experiments, we observed and compared the resulting stitching quality between the proposed method and existing methods. A representative seam finding algorithm [7] based on dynamic programming and a scheme depending on graph cut [8] and

TABLE 1: Experiment PC specification.

Processor	Intel® Core™ i7-4790 CPU @ 3.60 Ghz
Memory	16 GM
OS	Ubuntu 14.04 LTS 64 bits (VMware)

a Voronoi diagram method were chosen for comparison. We refer to the former as the seam finding and the latter as the graph cut.

We used four different sets of test images. Each set consists of two overlapping images each corresponding to left and right. Each image has the resolution of  $640 * 480$ . We used the computer with the specification of Table 1 for the experiments and implemented our algorithm by using an open source library, OpenCV 2.4.9 [12], with the parameters  $r$  and  $K$  set as 3 and 40, respectively.

Figure 5 shows the four sets of images which are used for the experiments. Each set consists of two images: left and right. The two images have common overlapping area by which they are stitched together. We select the test images such that overlapping areas are different features.

Figure 6 shows the stitching results of test image (a). Figures 6(a), 6(b), 6(c), and 6(d) are the results of the dynamic programming, the graph cut, the Voronoi diagram, and the proposed method, respectively. We can observe that, in all the methods except our method, the rail installed on the ceiling is not continuous, and a similar discrepancy is also found at the chair leg on the floor. However, the result of the proposed scheme is free from those flaws.

Figure 7 shows the stitching results of test image (b). Figures 7(a), 7(b), 7(c), and 7(d) are the results of the dynamic programming, the graph cut, the Voronoi diagram, and the proposed method, respectively. We can observe that, in all the methods except our method, the rectangular pillar in the middle has a very large mismatch. However, the result of the proposed scheme is free from those flaws.

Figure 8 shows the stitching results of test image (c). Figures 8(a), 8(b), 8(c), and 8(d) are the results of the dynamic programming, the graph cut, the Voronoi diagram, and the proposed method, respectively. We can observe that, in all



(a)



(b)



(c)



(d)

FIGURE 5: Four sets of test images. Each set consists of left and right images stitched to form a panorama.

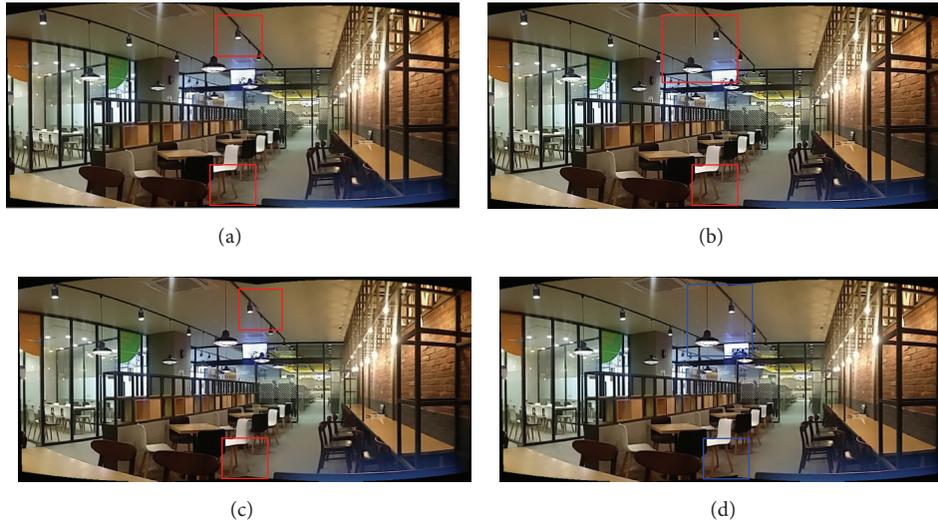


FIGURE 6: Result comparison for the test image (a) by the seam finding methods: (a) dynamic programming, (b) graph cut, (c) Voronoi diagram, and (d) proposed method.

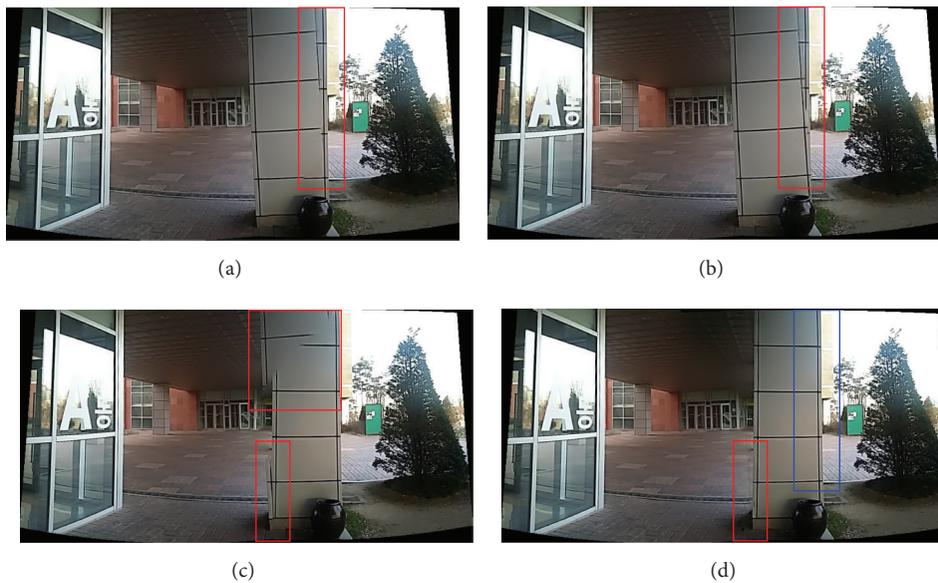


FIGURE 7: Result comparison the test image (b) by the seam finding methods: (a) dynamic programming, (b) graph cut, (c) Voronoi diagram, and (d) proposed method.

the methods except our method, each image has a similar discrepancy at the top middle part and also at the lower middle part. However, the result of the proposed scheme is free from those flaws.

Figure 9 shows the stitching results of test image (d). Figures 9(a), 9(b), 9(c), and 9(d) are the results of the dynamic programming, the graph cut, the Voronoi diagram, and the proposed method, respectively. We can observe that all the methods show similar results without noticeable flaws or mismatches.

Figure 10 shows the processing times of the methods. Given a set of left and right video streams as input to each method, we measured the times consumed for stitching each frame. It is observed that the elapsed times vary as frames proceed. It is because the contents of stitched images are different according to frames, resulting in different processing times. The Voronoi diagram was the fastest with average of 2.6 msec, while the graph cut was the slowest with average of 52.5 msec. The proposed method scored 36.1 msec which was better than the graph cut but was slower than the others.

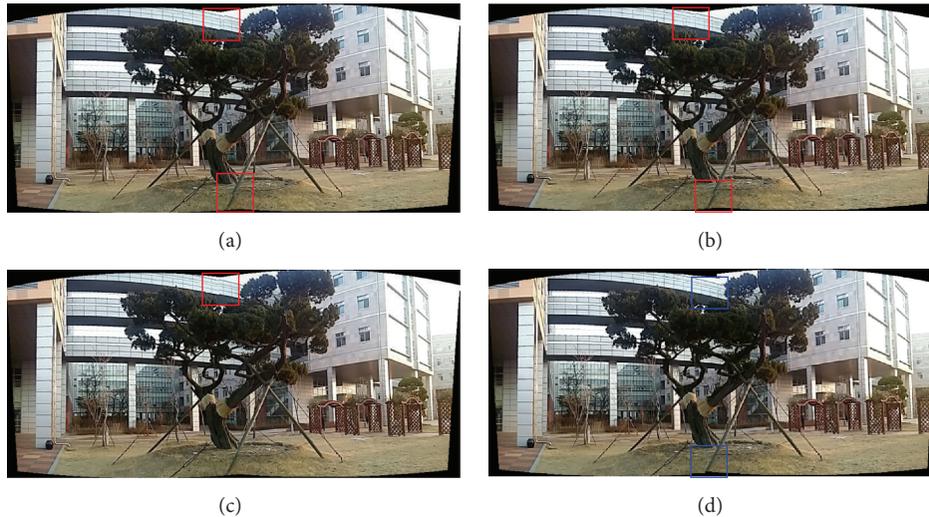


FIGURE 8: Result comparison for the test image (c) by the seam finding methods: (a) dynamic programming, (b) graph cut, (c) Voronoi diagram, and (d) proposed method.

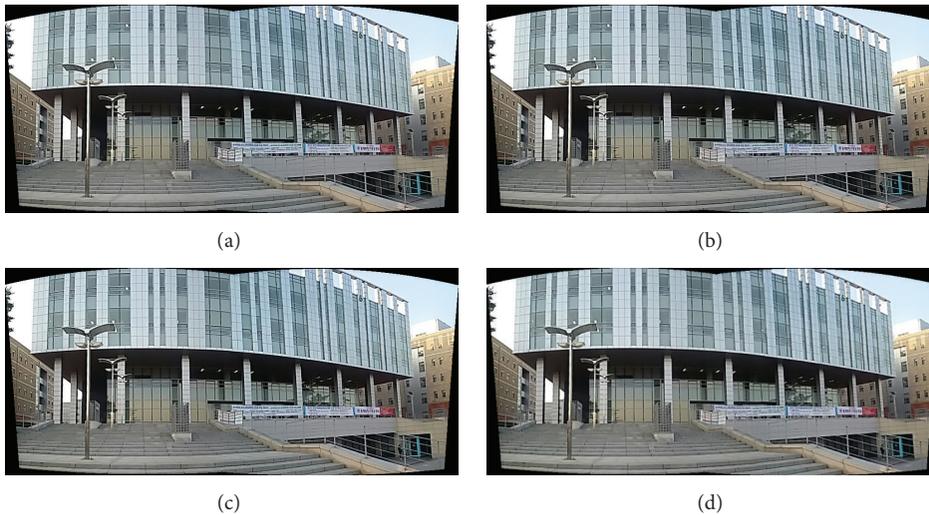


FIGURE 9: Result comparison for the test image (d) by the seam finding methods: (a) dynamic programming, (b) graph cut, (c) Voronoi diagram, and (d) proposed method.

However, considering that the resulting panorama quality of the proposed method is superior to the others, the increased time was spent for Gaussian filtering, entropy calculation to improve the results.

#### 4. Conclusions

We proposed an improved seam finding method which is important to smooth out the boundary of stitched images in panorama creation. For this, we defined a cost function to determine how each pixel affects stitching results. We were also able to reduce time consumption; instead of finding a seam over the whole area of overlapped region, we worked on a downscaled version of the overlapped image and interpolated to the original size. We carried out panorama creation

experiments in which our method as well as existing seam finding algorithms was used. Resulting panorama image quality was superior when our method was used, while the processing time of our method was comparable to those of others.

Before concluding the paper, it should be noted that the interpolation that was used to improve the performance of the proposed method can induce improper seam finding results, unless it is carefully tuned. Currently, we depended on algebraic interpolation. But the effects of such scheme will be evaluated and discussed in future research.

#### Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

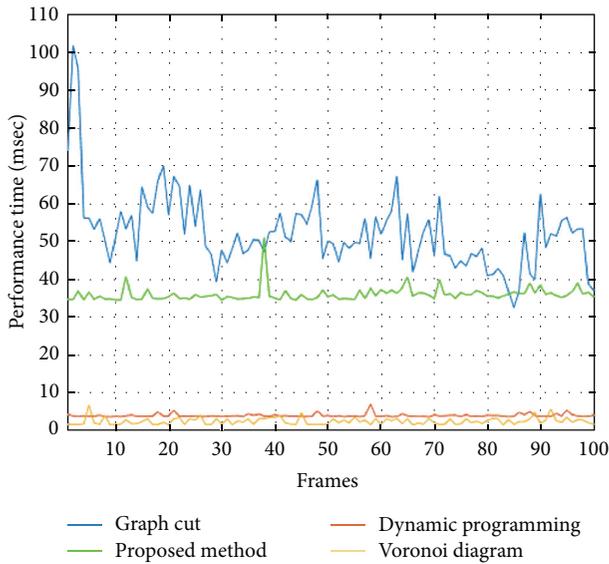


FIGURE 10: Performance comparison: elapsed times of the seam finding methods over video.

## Acknowledgments

This work was supported by the Incheon National University Research Grant in 2015.

## References

- [1] Panono, <https://www.panono.com/>.
- [2] Rotopan. <http://rotopan.com/>.
- [3] Pelco, <https://www.pelco.com/>.
- [4] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [5] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "SURF: speeded-up robust features," in *Proceedings of the European Conference on Computer Vision*, pp. 346–359, 2006.
- [6] S. Avidan and A. Shamir, "Seam carving for content-aware image resizing," *ACM Transactions on Graphics*, vol. 26, no. 3, pp. 1–9, 2007.
- [7] Y. Y. Boykov and M. P. Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images," in *Proceedings of International Conference on Computer Vision*, vol. 1, pp. 105–112, IEEE, Vancouver, Canada, 2001.
- [8] B. Kim, S. Lee, and N. Cho, "Real-time panorama image synthesis by fast camera pose estimation," in *Proceedings of the Signal & Information Processing Association Annual Summit and Conference (APSIPA ASC '12)*, pp. 3–6, Hollywood, Calif, USA, December 2012.
- [9] V. Kwatra, A. Schodl, and I. Essa, "Graphcut textures: image and video synthesis using graph cut," *ACM Transactions on Graphics*, vol. 22, no. 3, pp. 597–607, 2003.
- [10] N. Gracias, M. Mahoor, and S. Negahdaripour, "Fast image blending using watersheds and graph cuts," *Image and Vision Computing*, vol. 27, no. 5, pp. 597–607, 2009.
- [11] H. Gu, Y. Yu, and W. Sun, "A new optimal seam selection method for airborne image stitching," in *Proceedings of the International Workshop on Imaging Systems and Techniques*, pp. 159–163, Shenzhen, China, May 2009.
- [12] OpenCV. <http://opencv.org>.

## Research Article

# A Novel Encryption Algorithm Based on DWT and Multichaos Mapping

Wei Wang,<sup>1</sup> Haiyan Tan,<sup>1</sup> Yu Pang,<sup>1</sup> Zhangyong Li,<sup>1</sup> Peng Ran,<sup>1</sup> and Jun Wu<sup>2</sup>

<sup>1</sup>Chongqing University of Posts and Telecommunications, Chongqing 400065, China

<sup>2</sup>Chongqing Kaize Technology Co. Ltd., Chongqing 400042, China

Correspondence should be addressed to Yu Pang; pangyu@cqupt.edu.cn

Received 21 December 2015; Revised 17 March 2016; Accepted 28 March 2016

Academic Editor: Valerio Bellandi

Copyright © 2016 Wei Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Encryption of a digital image is very important especially in applications of body area networks (BANs) since the image may include a number of privacy. Past encryption methods have disadvantages of the small key space and low ability of resistance to attack. In this paper, we propose a new encryption algorithm based on discrete wavelet transform (DWT) and multichaos which has characteristics of the deterministic, pseudorandomness, and sensitivity of initial values. The image is first decomposed and spatial reconstructed by two-dimensional DWT and then is performed by multichaos matrices for space encryption. The experimental results indicate that the proposed algorithm has a large key space, high key sensitivity, and excellent ability of resistance to attack.

## 1. Introduction

In recent years with the rapid development of Internet and wireless sensor networks, data such as images and texts generally require encryption. For example, BANs play a significant role in remote medical monitoring which use many sensors surrounded with a human body [1]. The information is private and only utilized by authorized agencies, so encryption of the information is necessary. How to guarantee the safety of the multimedia information becomes a new direction of computer cryptography. Image network interaction is applied in many fields such as military cooperation, finance, and scene monitoring. Since images have some characteristics of large data, redundancy, correlation, and format consistency, image encryption has become a concerned topic [2].

Traditional image encryption techniques have caused concern. For instance, the methods of one-dimensional, two-dimensional, frequency domain, and hybrid encryption obtain abundant achievements. The technology of one-dimensional image encryption includes DES and AES as the representative of the modern cryptography system, S-box encryption, and SCAN language encryption, matrix transformation encryption, and DNA computation system [3–8]. These above spatial methods have advantages of fast calculation and easy implementation.

Chaos is a kind of unpredictable and similar random motion sensitive to initial values in a deterministic dynamical system. The pseudorandom sequence generated by chaotic map has good randomness, correlation, and complexity, which includes unique cryptography characteristics, so its derivative of the super chaotic map is widely used in the field of image encryption. The paper [9] is of significance in the literature to introduce a chaotic encryption algorithm with a diffusion sequence for applications of nonlinear functions. The contribution of the paper [10] is to implement the Baker chaotic map and sequence encryption. In past chaotic image encryption algorithms, the problem of subkey security exists in the papers [11–13] in the key analysis scheme of the paper [14] and the paper [15] adopts logistics chaotic encryption scheme with higher security than that of the papers.

The transform domain image encryption technology uses Fourier transform (FT), discrete cosine transform (DCT), and discrete wavelet transform (DWT) to perform conversion from spatial domain to transform domain and safely encrypt obtained coefficients; that is, the image is encrypted by changing the values or positions of the coefficients [16–20]. The paper [21] performs multilayer wavelet decomposition for the image and implements frequency domain scrambling in each block of every layer. This technique leads to the decrypted lossless image which becomes the key to be applied

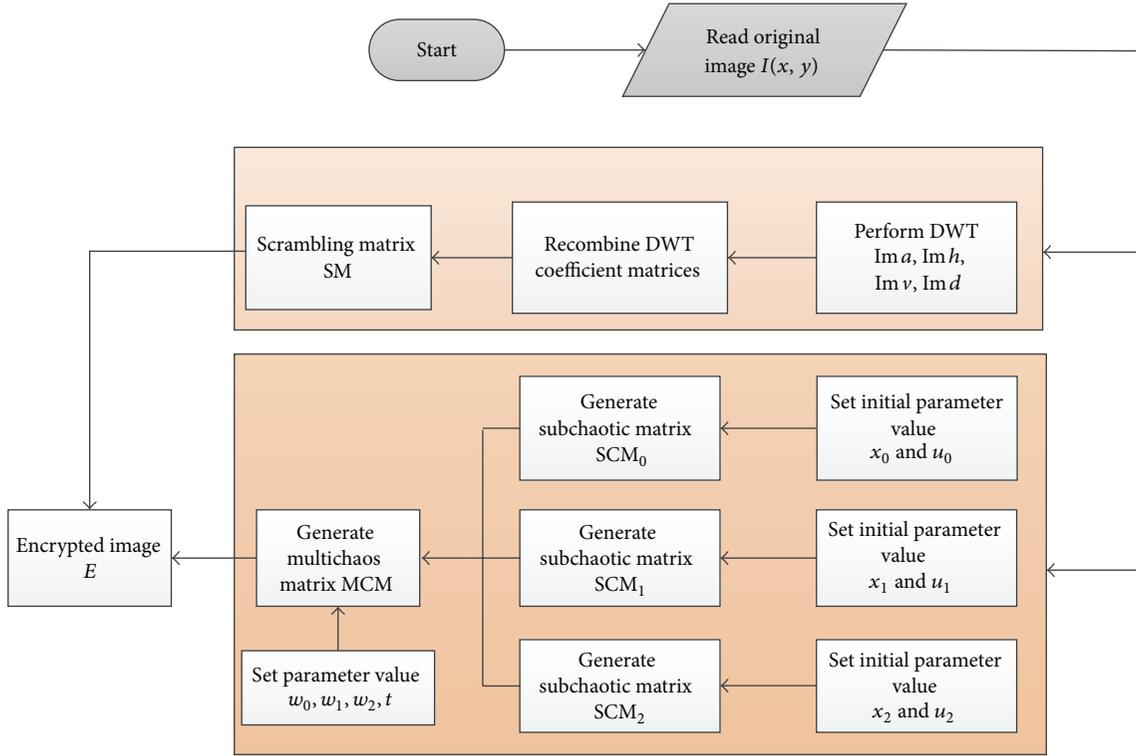


FIGURE 1: The flow of the proposed algorithm.

in frequency domain image encryption. The unique characteristics of chaos usually combine with the transform domain encryption mode to form mixed encryption algorithms with more advantages. Liu et al. [22] present a novel confusion and diffusion method, which generates a new key according to the original image and key and then uses piecewise linear chaotic map and Chebyshev chaotic map for DNA coding. Zhang et al. [7] code the original image to obtain a DNA sequence matrix, divide the matrix into several blocks, and finally use DNA computation combined with two logistic chaotic maps to achieve the goal of encryption.

In this paper, we take advantages of chaotic mapping and combine Haar wavelet to design the algorithm structure by a composite form of coefficient transform and multichaos. The proposed algorithm is sensitive to the initial state and system parameters with large key space and has low calculation complexity and high capability against attacks.

## 2. Theoretical Analysis of the Encryption Algorithm

**2.1. The Algorithm Flow of Image Encryption.** The steps of the algorithm are described as follows (as illustrated in Figure 1):

- (i) Read the original image.
- (ii) Perform DWT to obtain coefficient matrices  $Im\ a$ ,  $Im\ h$ ,  $Im\ v$ , and  $Im\ d$ .
- (iii) Recombine the four matrices to obtain the scrambling matrix  $SM$ .

- (iv) Set initial parameter values of  $x_0, x_1, x_2, u_0, u_1, u_2$  to generate three subchaotic matrices  $SCM_0, SCM_1$ , and  $SCM_2$ .
- (v) Set parameter values of  $w_0, w_1, w_2$ , and  $t$  combining the above three subchaotic matrices to generate the multichaos matrix  $MCM$ .
- (vi) Perform BitXOR operation of the scrambling image  $SM$  and multichaos matrix  $MCM$  to obtain the final encrypted image  $E$ .

**2.2. Multiscale Wavelet Decomposition in Two-Dimensional Images.** First, we transform the original image from time domain to wavelet domain using multiscale wavelet analysis, which is founded by the theory of function space. The multiscale decomposition of a two-dimensional gray image  $I(x, y)$  is shown as follows:

$$\begin{aligned}
 Im\ a &= \langle I(x, y), \phi(x - 2m, y - 2n) \rangle \\
 Im\ h &= \langle I(x, y), \Psi^1(x - 2m, y - 2n) \rangle \\
 Im\ v &= \langle I(x, y), \Psi^2(x - 2m, y - 2n) \rangle \\
 Im\ d &= \langle I(x, y), \Psi^3(x - 2m, y - 2n) \rangle.
 \end{aligned} \tag{1}$$

Here the Haar wavelet is used by one layer of decomposition to process the picture because the reconstruction has lossless feature. Different frequency coefficient matrices  $Im\ a, Im\ h, Im\ v$ , and  $Im\ d$  are obtained by the decomposition. The image size is  $M \times N$ ,  $\phi(x, y)$  is the two-dimensional

scale function, and  $\Psi(x, y)$  is the wavelet functions for corresponding positions. The initial encryption image  $E$  is obtained by recombining each layer coefficient matrix:

$$SM = \begin{bmatrix} \text{Im } a & \text{Im } h \\ \text{Im } v & \text{Im } d \end{bmatrix}. \quad (2)$$

Then we perform the multichaos operation for the initial encryption matrix  $E$ . The multichaos encryption matrix comes from subchaos matrices. The chaos mapping is the key of generating the subchaos matrices, which is a one-dimensional logistics mapping with characteristics of initial value sensitivity, parameter sensitivity, state ergodic property, and hybrid similarity stochastic:

$$l_n = \mu * l_{n-1} (1 - l_{n-1}). \quad (3)$$

$l_n$  represents the logistics chaotic mapping value in the pixel  $(i, j)$  by iteration. When the parameter  $\mu = [3.5699456, 4]$ ,  $l_0 \in (0, 1)$ , and  $n \in N$ ,  $l_n$  is in chaotic state. Given different initial values of  $l_0$  and  $\mu$ , we can obtain three chaotic matrices  $SCM_0$ ,  $SCM_1$ , and  $SCM_2$  generated by traversing each pixel to calculate the three different chaotic sequences  $\{l_n\}$ . Please note that the number of subchaotic matrices determines the time complexity of the algorithm. Using more subchaotic matrices may lead to higher encryption performance but the improvement is very limited by evaluation, so here only choosing three subchaotic matrices to form the multichaotic matrix is a good balance considering complexity and performance.

The multichaos encryption matrix MCM is calculated by the regulation parameters and three chaotic matrices:

$$\begin{aligned} \text{MCM}(i, j) = & W_0 (1 - t)^2 \text{SCM}_0(i, j) \\ & + 2W_1 (1 - t) \text{SCM}_1(i, j) \\ & + W_2 t^2 \text{SCM}_2(i, j). \end{aligned} \quad (4)$$

Here  $W_0$ ,  $W_1$ ,  $W_2$ ,  $t$  are matrix regulation parameters and all belong to the range  $(0, 1)$ . It is worth noting that they only participate in the generation of the multichaos matrix and so have little effect on the key space. Finally, the encrypted image  $E$  is obtained by

$$E = \text{MCM} \# \text{SM}. \quad (5)$$

Here the symbol “#” means the BitXOR operation.

The wavelet decomposition and reconstruction can effectively scramble the original image and largely cause pixel change compared to the Fourier transform and discrete cosine transform, while keeping the advantage of fast calculation because of only one layer performed. Here we use three logistic chaotic mapping with different initial conditions to generate three independent pseudorandom sequences leading to subchaotic matrices, so the created multichaotic matrix can resist the iterative attack from chaotic systems and have obvious advantages compared to classical chaos. Combining multiple chaotic systems can obtain more complex dynamic characteristics and become difficult to predict since

the application of the multichaotic matrices enhances the average changing intensity and sensitivity of initial parameter values. By mixing the wavelet transform and multichaos, the proposed method more excellently resists various attacks which can satisfy safety requirements of digital images and also the time complexity is acceptable.

### 3. Experimental Results and Analysis

To help investigate the security performance of the proposed algorithm, a program is developed and run on a 4 G memory, 3.2 GHz Intel(R) Core(TM) i5-4570 machine under Windows 8. The experimental data comes from the authoritative image database and we choose the image Lenna.jpg which can test each processing algorithm because it mixes detail, smooth area, shadow, and texture.

**3.1. Visual Results Display.** We compare the original image and encrypted image according to evaluation criteria [26]. The display results and the pixel histograms are shown, respectively, in Figures 2 and 3. The encrypted image in Figure 2(b) is completely different with the original and cannot be distinguished. The histogram of the original image exists in a narrow area (40–220) shown in Figure 3(a) and obviously has several maximum values in 50, 100, and 150. The histogram of the encrypted image in Figure 3(b) has wider distributed area and smooth values which means the characteristics of the image are covered up well.

**3.2. Analysis of Quantitative Results.** The most difference between image data and text data is that image data has strong correlation and a number of adjacent pixels have the same gray values or very small difference. If a data point and its adjacent data point still keep adjacent positions after scrambling, they are easy to be attacked by area analysis leading to low security. Adjacent elements include pixel in horizontal and vertical direction as well as diagonal direction. The correlation of adjacent elements is calculated as

$$\begin{aligned} \gamma &= \frac{\text{cov}(x, y)}{\sqrt{D(x)}\sqrt{D(y)}}, \\ \text{cov}(x, y) &= \frac{1}{N} \sum_{i=1}^N (x_i - E(x))(y_i - E(y)), \\ D(x) &= \frac{1}{N} \sum_{i=1}^N (x_i - E(x))^2, \\ E(x) &= \frac{1}{N} \sum_{i=1}^N x_i. \end{aligned} \quad (6)$$

Here  $x_i$  and  $y_i$  represent gray values of two adjacent pixels and  $N$  is the pixel number. In Table 1, the correlations in three directions are all beyond 0.9 that denotes that adjacent pixels have very similar gray values before encryption, while they decrease to under 0.01 after encryption, which means the much smaller correlation can make better resistance attack.

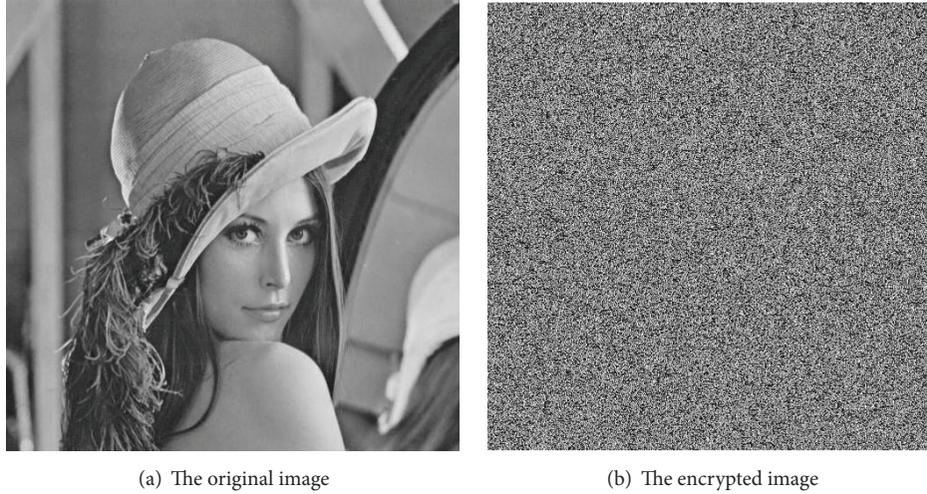


FIGURE 2: The encryption effect of the proposed algorithm.

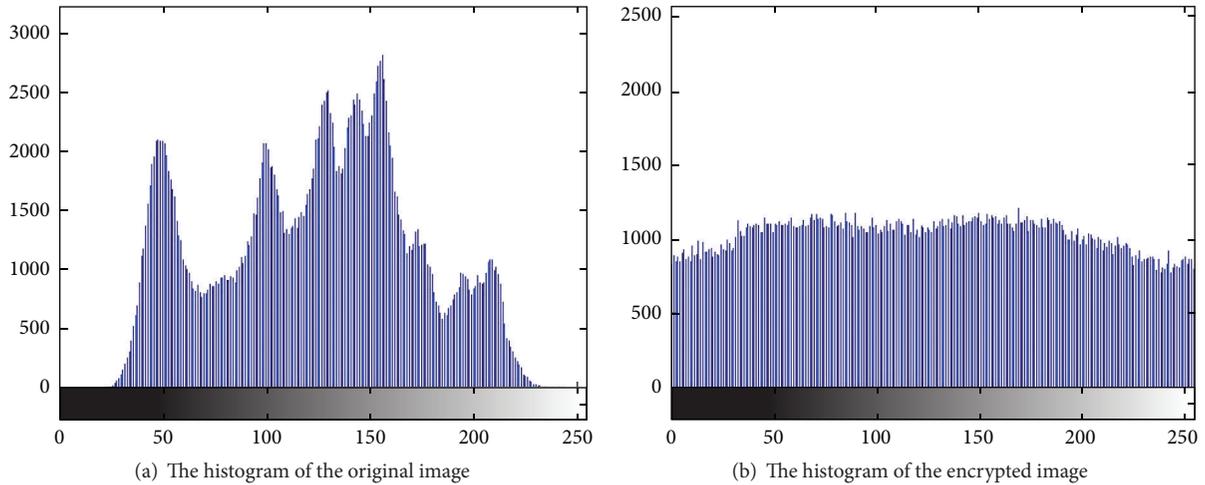


FIGURE 3: The original and encrypted statistical histograms.

TABLE 1: The correlation comparison before and after encryption of the image Lenna.jpg.

Correlation	Horizontal	Vertical	Diagonal
Before encryption	0.9388	0.9633	0.9417
After encryption	0.0005	-0.0003	0.0085

The goal of image encryption makes the encrypted image and original image as different as possible and hard to recognize. Obviously, the less fixed point ratio between two images represents more difference of the two images and better scrambling effect. The number of pixels change rate (NPCR) is defined as follows:

$$\text{NPCR} = \frac{\sum_{i=1}^M \sum_{j=1}^N D(i, j)}{MN} \times 100\%. \quad (7)$$

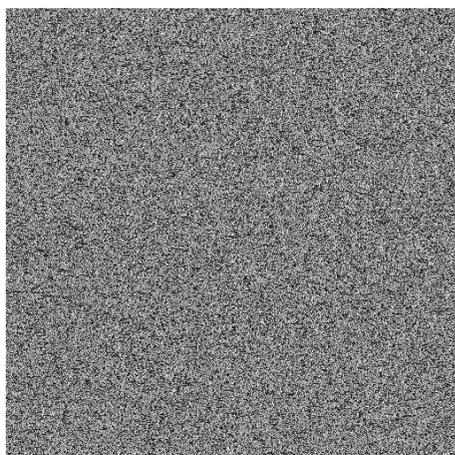
$D(i, j)$  is the gray value difference of the original and encrypted images in the pixel  $(i, j)$ .

Gray values of many pixels may change after encryption, so NPCR plays a good indicator to reflect the gray value change in number but noneffective to express the degree of gray value change. The average gray value change is necessary for evaluation. We give the definition of the unified average changing intensity (UACI):

$$\text{UACI} = \frac{1}{M \times N} \left[ \sum_{i,j} \frac{|I(i, j) - E(i, j)|}{255} \right] \times 100\%. \quad (8)$$

Here  $I(i, j)$  and  $E(i, j)$  represent the gray values of the original and encrypted images in the pixel  $(i, j)$ .

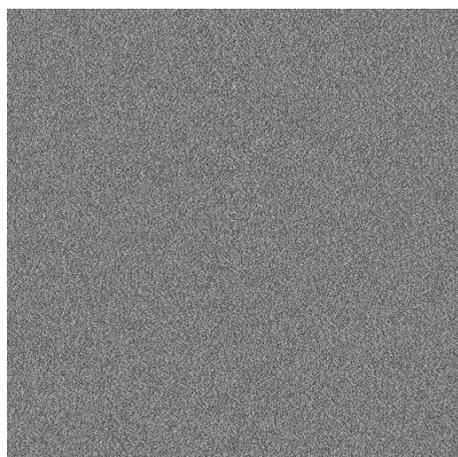
Furthermore, we use information entropy of the encrypted image to evaluate the degree of the average uncertainty. When the cipher has equal probability distribution leading to the maximum entropy value "8", the ideal random feature is achieved so the encrypted image has strong average



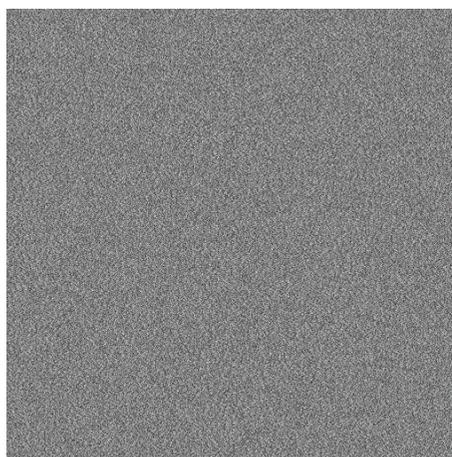
(a) The encrypted image



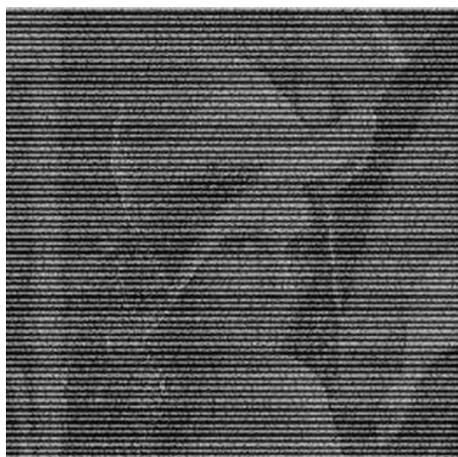
(b) The correct decrypted image



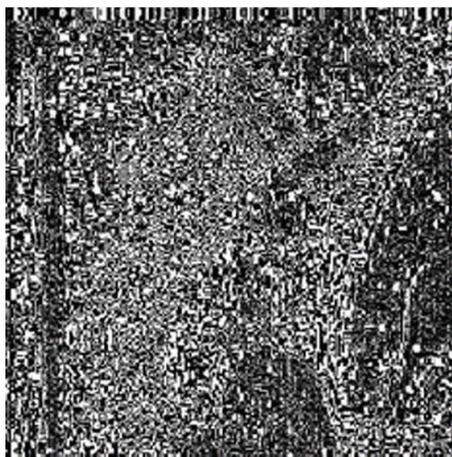
(c) The decrypted image with minor change of the parameter  $\mu$



(d) The decrypted image with minor change of key initial value  $x_0$



(e) The decrypted image with minor change of the parameter  $w$



(f) The decrypted image with minor change of the parameter  $t$

FIGURE 4: Test results of key sensitivity.

TABLE 2: The performance comparison of several algorithms.

Indicator	NPCR	UACI	$H$
Our algorithm	0.9987	0.3338	7.9989
The paper [23]	0.9962	0.3340	7.9992
The paper [24]	0.9985	0.3301	7.9904
The paper [25]	—	—	7.9822

uncertainty and high resistance to statistical attacks and entropy attacks. The entropy of each pixel is calculated as

$$H(M) = -\sum_{i=1}^{256} P(m(i)) \log_2 P(m(i)) = 7.9889 \approx 8. \quad (9)$$

Here  $P(m(i))$  represents the probability that each gray value shows.

We compare the proposed algorithm to other papers in Table 2 which all use methods of chaotic encryption. The paper [25] does not provide indicators of NPCR and UACI. Our algorithm can obtain the maximum NPCR value and the paper [23] plays best in UACI and  $H$ , so each of the two algorithms has its own advantages and both have better performance than the others.

**3.3. Analysis of Security Key.** The composition and size of the key space determine the security of encryption algorithms [26–28]. The proposed algorithm consists of two stages: the first stage combines two-dimensional DWT to scramble the image pixels and the second finishes pixels diffusion using multichaos. Therefore, the algorithm has various key combination forms with the parts of wavelet parameters and multichaos parameters.

We test the correct key and approximate key in the experiment of key sensitivity. The results are shown in Figure 4.

Figure 4(a) is the encrypted image and Figure 4(b) is the correct decryption image. We change parameters in the key with a very minor value ( $10^{-8}$ ) to test sensitivity of initial values. The results in Figures 4(c) and 4(d) denote that the decrypted images cannot be distinguished and have no relation with the original image when the parameters have slight alternations. Figures 4(e) and 4(f) show the decrypted images when the parameters  $w$  and  $t$  are changed slightly. Obviously, these two images still maintain some features of the original image and the performance of resistance to attack is much worse than that of Figures 4(c) and 4(d). The reason is that the parameters of  $\mu$  and  $x_0$  need multiple iterations to form the subchaotic matrices SCM according to (3), so the slight alternation may be propagated and enlarged leading to totally confused and unpredictable results. On the contrary, the parameters  $w$  and  $t$  only occur in (4) to form the final multichaotic matrix, so the small change has relatively minor effect.

According to Figure 4, very small change of parameters  $\mu$  and  $x_0$  cause completely different results. Because of almost unlimited number of  $\mu$  and  $x_0$  in their ranges and considering that each subchaotic matrix has its own parameters, the key space beyond  $10^{80}$  is much larger than that of classical chaos

algorithms, so the proposed algorithm has stronger capability to resist exhaustive attack.

## 4. Conclusions

The proposed algorithm in this paper encrypts image pixel values and pixel locations according to basic idea of image encryption. Pixel gray change is processed by multichaos and pixel scrambling uses DWT for pixel encryption, which changes the traditional encryption idea of image resolution invariability.

The algorithm effectively hides the image size and improves the safety with advantages of large key space, sensitive to key and clear text, and easy to fast implementation. The compound chaotic system makes attackers difficult to analyze and estimate leading to high system complexity and safety. Furthermore, the algorithm has strong expansibility. Using other chaotic systems or external generators to create a key can enhance more safety of the algorithm. Dynamically setting the length of the key and packet according to the parameter adjustment is helpful to broaden serviceability of the algorithm and suitable to applications of secure communications and network security.

## Competing Interests

The authors declare that they have no competing interests.

## Acknowledgments

This project is supported by the National Science Foundation of China (Grant nos. 61471075 and 61301124), Chongqing Integrated Demonstration Project (CSTC2013jcsf10029), Chongqing Talented Youth Development Plan (cstc2013kjrc-qnc10001), Wenfeng Innovation Foundation of CQUPT, 2013 University Innovation Team Construction Plan Funding Project of Chongqing (Smart Medical System and Key Techniques), and Chongqing Key Laboratory Improvement Plan (Chongqing Key Laboratory of Photoelectronic Information Sensing and Transmitting Technology).

## References

- [1] M. Patel and J. Wang, "Applications, challenges, and prospective in emerging body area networking technologies," *IEEE Wireless Communications*, vol. 17, no. 1, pp. 80–88, 2010.
- [2] M. Xiang-Feng, C. Lu-Zhong, and Y. Xiu-Lun, "Information security system by iterative multiple-phase retrieval and pixel random permutation," *Applied Optics*, vol. 45, no. 14, pp. 3289–3297, 2006.
- [3] V. M. El-Zoghdy, Y. A. Nada, and A. A. Abdo, "How good is the DES algorithm in image ciphering," *International Journal of Advanced Networking and Applications*, vol. 2, no. 5, pp. 796–803, 2011.
- [4] B. Subramanyan, V. M. Chhabria, and T. G. S. Babu, "Image encryption based on AES key expansion," in *Proceedings of the 2nd International Conference on Emerging Applications of Information Technology (EAIT '11)*, pp. 217–220, Kolkata, India, February 2011.

- [5] R.-J. Chen and S.-J. Horng, "Novel SCAN-CA-based image security system using SCAN and 2-D von Neumann cellular automata," *Signal Processing: Image Communication*, vol. 25, no. 6, pp. 413–426, 2010.
- [6] V. P. Singh, R. Beg, and B. Mishra, "Practical approach on image encryption and decryption technique using matrix transformation," *MIT International Journal of Computer Science & Information Technology*, vol. 3, no. 1, pp. 38–41, 2013.
- [7] Q. Zhang, L. Guo, and X. P. Wei, "Image encryption using DNA addition combining with chaotic maps," *Mathematical and Computer Modelling*, vol. 52, no. 11-12, pp. 2028–2035, 2010.
- [8] H. Liu, X. Wang, and A. Kadir, "Image encryption using DNA complementary rule and chaotic maps," *Applied Soft Computing Journal*, vol. 12, no. 5, pp. 1457–1466, 2012.
- [9] J. S. A. E. Fouda, J. Y. Effa, S. L. Sabat, and M. Ali, "A fast chaotic block cipher for image encryption," *Communications in Nonlinear Science & Numerical Simulation*, vol. 19, no. 3, pp. 578–588, 2014.
- [10] E. Solak, C. Cokal, O. T. Yildiz, and T. Biyikoğlu, "Cryptanalysis of Fridrich's chaotic image encryption," *International Journal of Bifurcation and Chaos in Applied Sciences and Engineering*, vol. 20, no. 5, pp. 1405–1413, 2010.
- [11] Y. Y. Sun, R. O. Kong et al., "An image encryption algorithm utilizing Mandel bot set," in *Proceedings of the International Workshop on Chaos-Fractal Theory and Its Applications*, pp. 170–173, 2010.
- [12] C.-C. Chang and T.-X. Yu, "Cryptanalysis of an encryption scheme for binary images," *Pattern Recognition Letters*, vol. 23, no. 14, pp. 1847–1852, 2002.
- [13] X. J. Tong and M. R. Cui, "Image encryption with compound chaotic sequence cipher shifting dynamically," *Image and Vision Computing*, vol. 26, no. 6, pp. 843–850, 2008.
- [14] C. Q. Li, S. J. Li, M. Asim, J. Nunez, G. Alvarez, and G. Chen, "On the security defects of an image encryption scheme," *Image and Vision Computing*, vol. 27, no. 9, pp. 1371–1381, 2009.
- [15] N. K. Pareek, V. Patidar, and K. K. Sud, "Image encryption using chaotic logistic map," *Image and Vision Computing*, vol. 24, no. 9, pp. 926–934, 2006.
- [16] M. Mishra and P. Mishra, "Image encryption using fibonacci-lucas transformation," *International Journal on Cryptography and Information Security*, vol. 2, no. 3, pp. 131–141, 2012.
- [17] A. A. Abd El-Latif, X. Niu, and M. Amin, "A new image cipher in time and frequency domains," *Optics Communications*, vol. 285, no. 21-22, pp. 4241–4251, 2012.
- [18] S. Sasidharan and D. S. Philip, "A fast partial image encryption scheme with wavelet transform and RC4," *International Journal of Advances in Engineering & Technology*, vol. 1, no. 4, pp. 322–331, 2011.
- [19] A. Nikolaidis, "Asymptotically optimal detection for additive watermarking in the DCT and DCT Domian," *IEEE Transactions on Image Processing*, vol. 12, no. 5, pp. 563–571, 2003.
- [20] Y. Liang, G. Liu, N. Zhou, and J. Wu, "Image encryption combining multiple generating sequences controlled fractional DCT with dependent scrambling and diffusion," *Journal of Modern Optics*, vol. 62, no. 4, pp. 251–264, 2015.
- [21] G. Ye, "A block image encryption algorithm based on wave transmission and chaotic systems," *Nonlinear Dynamics*, vol. 75, no. 3, pp. 417–427, 2014.
- [22] H. J. Liu, X. Y. Wang, and A. Kadir, "Image encryption using DNA complementary rule and chaotic maps," *Applied Soft Computing*, vol. 12, no. 5, pp. 1457–1466, 2012.
- [23] J. S. Armand Eyebe Fouda, J. Y. Effa, S. L. Sabat, and M. Ali, "A fast chaotic block cipher for image encryption," *Communications in Nonlinear Science and Numerical Simulation*, vol. 19, no. 3, pp. 578–588, 2014.
- [24] A. Jolfaei and A. Mirghadri, "An image encryption approach using chaos and stream cipher," *Journal of Theoretical and Applied Information Technology*, vol. 19, no. 2, pp. 117–125, 2010.
- [25] A. H. Abdullah, R. Enayatifar, and M. Lee, "A hybrid genetic algorithm and chaotic function model for image encryption," *AEU-International Journal of Electronics and Communications*, vol. 66, no. 10, pp. 806–816, 2012.
- [26] J. Ahmad and F. Ahmed, "Efficiency analysis and security evaluation of image encryption schemes," *International Journal of Video & Image Processing and Network Security*, vol. 12, no. 4, pp. 18–31, 2012.
- [27] M. Su, W. Wen, and Y. Zhang, "Security evaluation of bilateral-diffusion based image encryption algorithm," *Nonlinear Dynamics*, vol. 77, no. 1-2, pp. 243–246, 2014.
- [28] A. K. Selvi and D. M. M. Sathik, "Efficiency analysis and security evaluation of block based image encryption schemes," *Digital Image Processing*, vol. 7, no. 1, 2015.

## Research Article

# Context-Aware Mobile Sensors for Sensing Discrete Events in Smart Environment

**Awais Ahmad,<sup>1</sup> M. Mazhar Rathore,<sup>1</sup> Anand Paul,<sup>1</sup> Won-Hwa Hong,<sup>2</sup> and HyunCheol Seo<sup>2</sup>**

<sup>1</sup>*School of Computer Science and Engineering, Kyungpook National University, 80 Daehak-ro, Buk-gu, Daegu 702-701, Republic of Korea*

<sup>2</sup>*School of Architecture, Civil, Environmental and Energy Engineering, Kyungpook National University, 80 Daehak-ro, Buk-gu, Daegu 702-701, Republic of Korea*

Correspondence should be addressed to Anand Paul; [paul.editor@gmail.com](mailto:paul.editor@gmail.com) and Won-Hwa Hong; [hongwh@knu.ac.kr](mailto:hongwh@knu.ac.kr)

Received 25 December 2015; Accepted 28 March 2016

Academic Editor: Marco Anisetti

Copyright © 2016 Awais Ahmad et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Over the last few decades, several advancements in the field of smart environment gained importance, so the experts can analyze ideas for smart building based on embedded systems to minimize the expense and energy conservation. Therefore, propelling the concept of smart home toward smart building, several challenges of power, communication, and sensors' connectivity can be seen. Such challenges distort the interconnectivity between different technologies, such as Bluetooth and ZigBee, making it possible to provide the continuous connectivity among different objects such as sensors, actuators, home appliances, and cell phones. Therefore, this paper presents the concept of smart building based on embedded systems that enhance low power mobile sensors for sensing discrete events in embedded systems. The proposed scheme comprises system architecture that welcomes all the mobile sensors to communicate with each other using a single platform service. The proposed system enhances the concept of smart building in three stages (i.e., visualization, data analysis, and application). For low power mobile sensors, we propose a communication model, which provides a common medium for communication. Finally, the results show that the proposed system architecture efficiently processes, analyzes, and integrates different datasets efficiently and triggers actions to provide safety measurements for the elderly, patients, and others.

## 1. Introduction

Recently, the concept of smart building is developed and researchers and other experts are trying to provide the best services to the people who live in such buildings. Since it is difficult to provide their best facilities in a real-time environment due to high cost of the home appliances [1], many researchers are using simulation environment that decreased the cost of test, design, and environment [2, 3]. With a view to design smart building, the rapid growth of mobile sensors is required to be connected with each other with the help of the Internet that creates IoT. Such network generates enormous amount of data, usually termed as Big Data. Similarly, enabling seamless connectivity with existing networks and proactive operation based on different factors (context-aware computation) are mandatory in IoT. The goal of IoT is to let the computer identify information without human interaction. However, it depends on three

factors, such as (1) understanding users and its appliances, (2) architecture and communication, and (3) analytical tools to support smart behavior. IoT and cloud computing can be considered as two major technologies, which have been developed in order to extend embedded computing. The embedded computing integrates different technologies with the daily life. Similarly, human-centric embedded computing focuses on the specific domain by means of whom should be benefited. The devices connected wirelessly have the ability to act as embedded sensors using MEMS and wireless communication. These devices are known as nodes in a mobile wireless sensor network (MWSN). Cloud computing uses the Internet to provide scalable, reliable ubiquitous computing by acting as a receiver of data from ubiquitous sensors.

IoT has been defined in multiple ways by multiple interest groups as IoT is a combination of middleware, sensors, and knowledge [4]. However, according to Radio-Frequency

Identification (RFID) group, IoT can be defined as a network in which objects are accessible uniquely via standard protocols [5]. In addition, things are associated with each other with the help of IoT that is able to sense and share information without human intervention. Hence, the researchers from various fields describe IoT as the interconnection of devices that share information among different platforms via standard platforms. Such facilities can be achieved via ubiquitous sensing, data analytics, and so forth in which various entities are involved that are RFID, WSN, their addressing schemes, data storage, and analytics.

To make IoT more appealing, a traditional application, that is, smart building, is considered where embedded devices (such as sensors and actuators) are considered autonomously and is remotely controlled by means of the Internet. The mentioned technology is used to empower monitoring application with the security features. The embedded devices are used to sense user activities and control electric appliances according to the information collected by these devices. Such development leads us toward the concept of smart building where the information is being processed, analyzed, and predicted or gives response to the outer stimuli based on the received information according to certain rules. The rules that are triggered by actions can be uttered as the form of condition action. For instance, an action can be defined as “working in an office in a bright light, and all the lights are switched on, the lights will become dim automatically.” In designing smart building systems, rules are important components that provide flexible control. The basic idea of these rules is to distinguish between logic and data, which helps in making maintainable parts independent. Therefore, in the literature, many rule engines are designed that are used to reduce the cost of designing, developing, and delivering software [6–9]. For instance, in a traditional wireless sensor and actuator network, different kinds of embedded sensors are deployed that are used to collect environmental data. Each device is equipped with actuator capable of receiving control commands. Consequently, results are contributing a lot of rule sets. These rules are generally quite complex to handle huge volume of big data; this cannot be directly applied to smart building systems.

Having understood the feasibility and potential of embedded devices, in this paper, we propel the concept of smart home to a further extent and introduce a notion of context-aware low power mobile sensors for sensing discrete events in embedded systems. In the proposed system, mobile sensors are deployed in a building that senses discrete events (such as temperature, user activities, and body area network for healthcare) and share this information using the same communication medium, that is, the Internet. The shared medium is supported by the Internet where a variety of devices are interconnected with each other. The nature of these devices is heterogeneous, which requires a unique platform to exchange useful information. Moreover, in order to achieve this, an architecture is also proposed in which an efficient communication among various embedded devices takes place in a smart building based on the contextual information. It can be viewed that these devices (such as smart watch, healthcare, Kinect Xbox 360, Internet of Vehicles,

and GPS) continuously monitor the physical entities, and, when required, automatic or controlled physical system gives alert to the specific event to improve healthcare, security, accidents, fire brigade system, and so on. This system is connected to the Internet with the help of Wi-Fi (IEEE 802.11) and the third generation (3G) of mobile telephone.

The remainder of this paper is organized as follows. Section 2 presents a detailed description of the background and related studies. In Section 3, we described some of the IoT characteristics. Section 4 presents the proposed system architecture for context-aware low power mobile sensors for sensing discrete events. In Section 5, a detailed analytical and simulation analysis and a conclusion of the paper are presented.

## 2. Background and Related Studies

In this section, we provide the background of WoT that could be integrated with the IoT along with the related studies in the field of IoT and smart home. Much of the research work has been done in the field of smart buildings and smart home. However, still the short holes are there to fill the gap. Therefore, the researchers are still working to make building smarter in order to take a smart and intelligent decision in the case of any emergency and disaster.

Asimakopoulou and Bessis [10] work in sensing of building and crowd in order to handle disasters and serious problems occurring in the building. Do disaster management. They took the concept of crowdsourcing to utilize it in smart buildings and cities in order to effectively manage the disasters. Some other research LABs and organizations tried to build a practical testbed of the smart building in order to measure and sense the internal environment of the building. A testbed place in the University of Southern California [11] is selected to build because of its diverse nature of mix spaces and classroom usage and scheduling. The testbed was facilitated by advanced indoor sensors actuators and intelligent building management systems, suitable for broad range of research activities. The complete communication infrastructure is built to manage the whole transmission between sensors and a centralized server. The centralized server controls the actuators by sending signals to them. The number of sensors is deployed to get the real data related to building environment and other things such as lighting conditions, airflow, temperature, and CO<sub>2</sub> and other toxic gases. Similarly, Lee et al. [12] and Ahmad et al. [13, 14] also design a smart building and smart society by focusing on the energy aspect while considering smart grids in building. They investigated getting and controlling data by accessing the energy resources in a smart building through smart grid infrastructure. However, they did not focus on the energy service interface (ESI), which is considered as the main communication gateway for a smart building. They provide the illustration of interactive energy services between the grid infrastructure and smart buildings.

Some of the authors also provide the conceptual simulation of smart building and smart homes. Bidhandi and his research colleagues [15] divided their simulation process into two phases. In the initial phase, the person can make

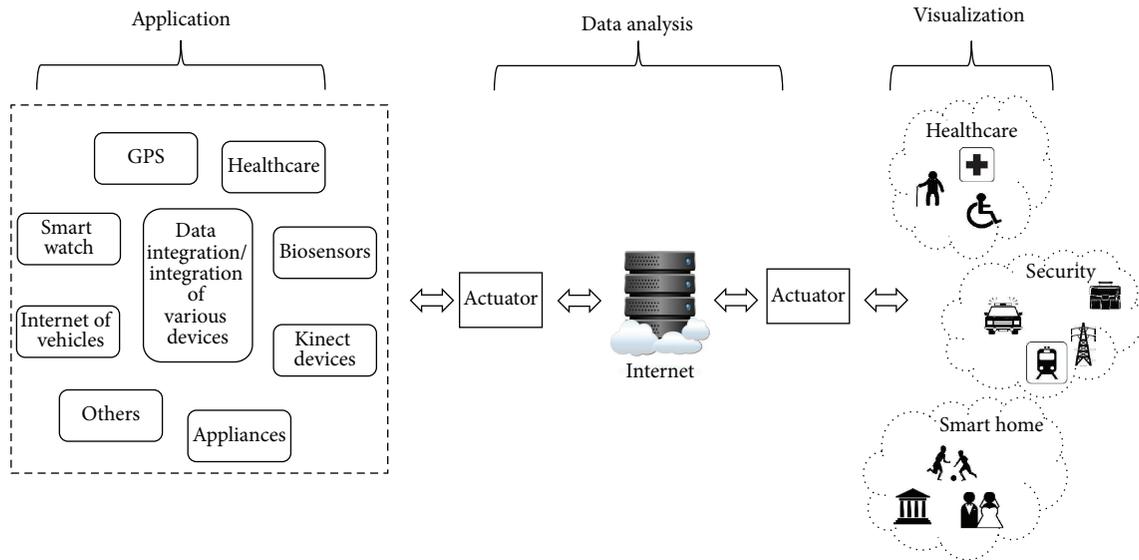


FIGURE 1: Overview of designed system.

design planning their home floor or building floor manually or through CAD software. Later their design allows adding home appliances and equipment to the building design. In the next phase, that is, phase two, they provide the simulation by two methods while defining scenario and analyzing the building reaction on various scenarios. They claimed to increase the features of flexibility and usability increased for the user to easily simulate any type of smart building with various features and ideas by reducing the overall costs. Moreover, in [16], authors developed new simulator based on the context information of smart homes by focusing on the main goal of context-aware system design. In [3, 17–19] also is presented a multipurpose simulator based on the scenario for the smart house atmosphere which again is an object oriented one and is capable of defining various sensors and placing them on the floor plan of the building. The user is also able to sketch the plan of the house. In [18], a configurable context-aware simulator for smart home systems is presented, in which the open service gateway initiative is considered for being operated in the framework of the smart house. And with such framework, we can make sure of the completion over time and also a connection to the outer world. Remote controlling, trouble diagnosis, and management are supported too. In [20], there is a full, conclusive survey about the architecture of service oriented simulator for smart house management. That is a service oriented perspective of making and using a simulator founded on reality and also a perspective of virtual integration with devices in a building model.

In rule based system in the smart building, broad work has been done, especially in processing schemes [21, 22]. In the terms of the rule engine, they mainly consist of two aspects: the RETE algorithms and the complex event processing mechanism. RETE is a customary algorithm for the rule engine developed by Forgy [23, 24] and first engaged in production systems and more M2M and V2V event engines shall be considered later [25–27].

### 3. Context-Aware Low Power Based Mobile Sensing System

In this section, we discuss the breakdown of the proposed system into two parts, that is, communication model and the architecture for context-aware low power mobile sensors for sensing discrete events in embedded systems. Before propelling toward the architecture for context-aware low power mobile sensors for sensing discrete events in embedded systems, it is worth presenting the overview of the proposed architecture.

*3.1. Overview.* The proposed context-aware low power mobile sensors for sensing discrete events in embedded systems consists of three major units, such as application unit, data analysis unit, and visualization unit, as shown in Figure 1. The application unit is comprised of the healthcare system, home security system, and smart building system. In the healthcare system, we have employed source ID/location-based 6LoWPAN wireless body area network mobility management scheme as shown in Figure 2. In the architecture, a group of 6LoWPAN sensors are considered, which are attached to human body. In this group of sensors, there is one coordinator that is used for exchanging control messages with Primary Mobile Device (PMD). Each PMD and sensor have a 128-bit global unique device identifier (GDID) [28]. GDID establishes end to end communication provided with the link-layer addresses as the Access Identifier (AID). The GDID has the information about its home domain. However, Local LOCators (LLOC) and Global LOCators (GLOC) identify location of PMDs within the home domain. Since the GLOC denotes the IP address of Access Gateway (AGW) that is used for interdomain communication. In the given scenario, each AGW contains home GDID and visitor GDID register. HGR contains the mapping information about the GDID-LOC and VGR contains GDID-LLOC mapping information for the visited PMDs.

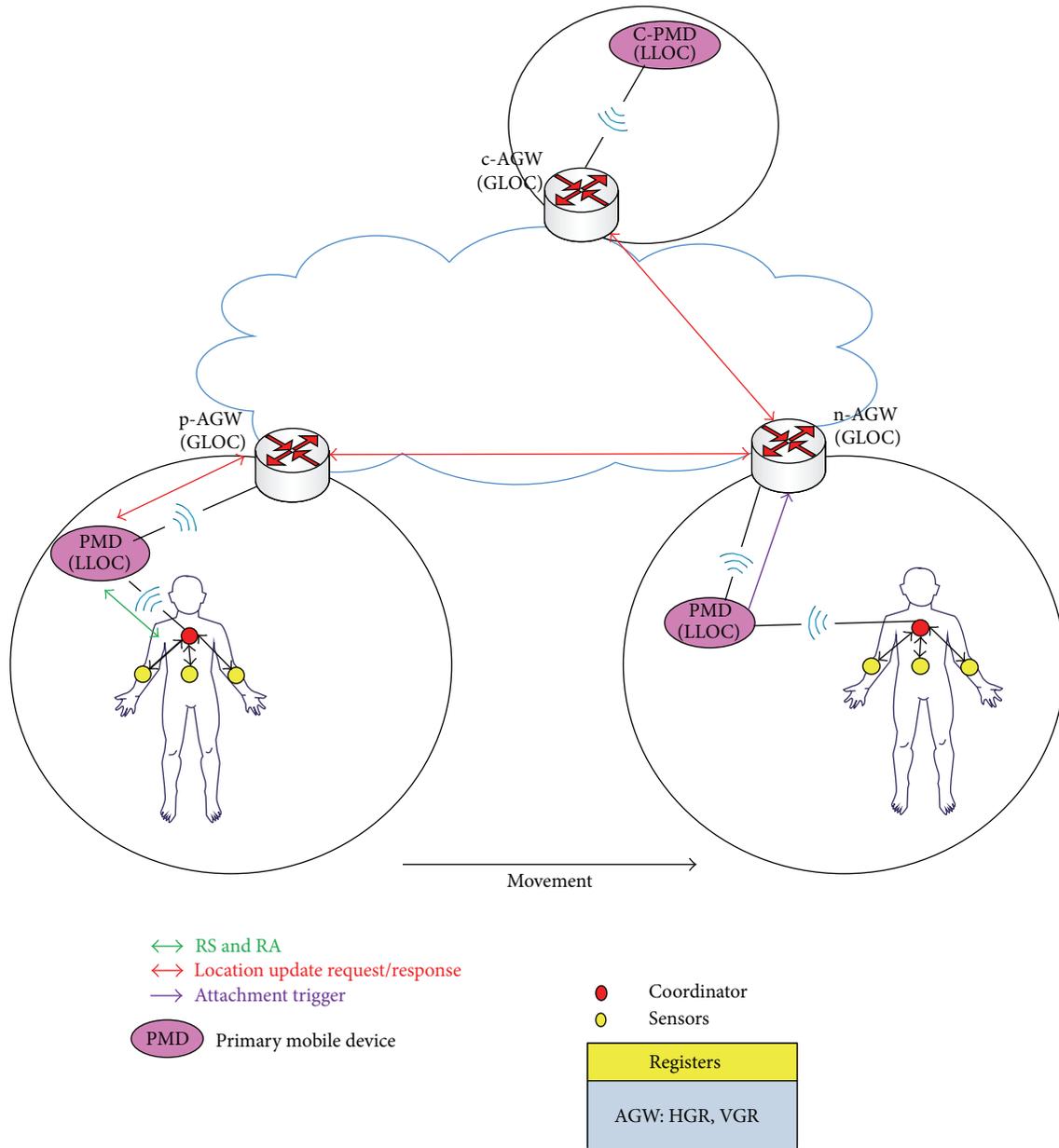


FIGURE 2: Network model for mobile body area network.

In the proposed architecture, only one-time Router Solicitation (RS) and Router Advertisement (RA) messages are sent by the coordinator and thus reduce lots of control messages. Initially, the PMD communicates with correspondent PMD (C-PMD) in the previous AGW (p-AGW) domain. Now, the PMD moves to a new AGW (n-AGW). Finally, the proposed smart building is equipped with the electronic appliances, which are interconnected with each other with the help of Bluetooth and ZigBee technologies.

3.2. *Communication Model.* We propose a couple of different technologies for communications in the smart home. The ZigBee technologies are used to cover the entire smart home because the ZigBee technology is considered as better than IEEE 802.11 in providing LAN services and advantages of

ZigBee include low power and low duty embedded systems. Hence, in the same way, we use Bluetooth Low Energy (BLE) device inside a smart home for covering an entire room. The actual purpose of using BLE is that it works efficiently for Personal Area Network (PAN). Therefore, we present an overview of the communication model of both technologies in the following subsections.

3.3. *Bluetooth Low Energy (BLE).* Bluetooth Low Energy (BLE) is considered one of the latest versions of the Bluetooth v4.0, which is designed for the low power consumption devices, such as sensors and wearable devices. The Bluetooth Low Energy (BLE) consists of two layers, that is, upper and lower. The upper layer provides the functionality of error control and flow control and lower layers handle the

transmission of bits over the physical medium. The upper layer is again divided into three types, that is, Logical Link Control and Adaptation Protocol (L2CAP), Generic Access Profile (GAP), and Generic Attribute Protocol (GAP). The Logical Link Control and Adaptation Protocol (L2CAP) is used for fragmentation and reassembling of large packets and the multiplexing of the data from upper layers. The Bluetooth Low Energy (BLE) is using Adaptive Frequency Hopping Spread spectrum (AFHSS), which is similar to classical Bluetooth. However, the BLE uses 40 channels, each of 2 MHz, bit rate of 1 Mbit/s, and transmit power equal to 10 mW. As BLE uses a bit rate of 1 MHz, it is unable to provide the voice capability. Because of these features, the BLE is considered in low energy devices such as mobile phones and wearable devices.

The BLE devours 90% (0.01 to 0.05 W) of the vitality on the advertising and examining process, though we talk about the advertising and examining process in detail. The BLE works in three distinct ways, that is, notice of the BLE devices, filtering of the accessible devices, and starting the connection, not at all like the classical Bluetooth. The BLE has three unique channels for advertising, that is, 37, 38, and 39. The BLE has 40 channels out of which three channels are utilized for the promotion purpose. Thus, the BLE continues sending the messages utilizing the ADV\_IND packet data information units over the three channels during the event. An arbitrary measure of deferral is utilized between the notices of the channels to avoid the crash between two advertisers. The BLE standard characterizes the advertisement interval for each of the three channels, and the notice ought to be somewhere around 20 and 10.24 ms having a number different of 0.625 ms. Likewise, the delay should be under 10 ms [29]. Once the advertisement stage is over, BLE will listen to the reactions from the accessible devices on the same channels. Along these lines, the gadget will go into filtering mode after add stage. The scanner responds to the advertisement message. The standard characterizes window time of 10.24 s for filtering the accessible gadgets. Keeping these attributes of BLE innovation, we use it in the proposed smart building design.

**3.4. ZigBee Protocol.** The ZigBee protocol is based on the IEEE 802.15.4 standard and can be used to transfer data over a distance of 10 to 100 meters. In addition, ZigBee is considered for low power devices since it requires less energy for their communication. Correspondingly, ZigBee consumes more energy on the transmission of data. Similar to BLE, ZigBee also operates on 2.5 GHz, that is, unlicensed band worldwide, which uses 16 channels with a 5 MHz space and 2 MHz bandwidth. The ZigBee uses direct-sequence spread spectrum (DSSS) coding technique. Moreover, there are two types of communication modes that are present for data transmission, that is, beacon-enabled and non-beacon-enabled. In the case of beacon-enabled approach, the network coordinator periodically broadcasts the beacon message in the PAN Network. The devices in the PAN network synchronized with the coordinator upon receiving the beacon messages. In the case of the non-beacon-enabled communications, randomly broadcast the beacon messages. If a device wants to send

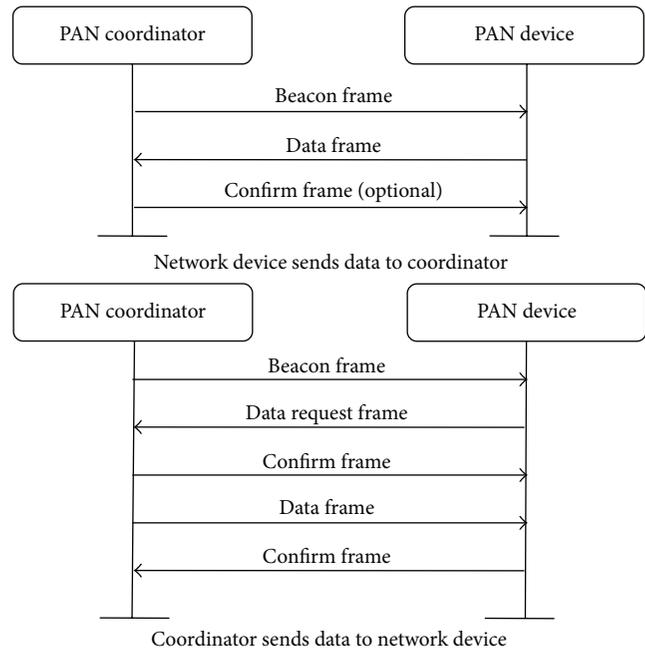


FIGURE 3: Beacon-enabled communications in ZigBee protocol.

data it waits for a random amount of time and then senses (CSMA-CA) the channel; if the channel is available, it starts the data sending; otherwise it switches to waiting state. The framework of the beacon-enabled and non-beacon-enabled communication is shown in Figures 3 and 4, respectively [30]. In the case of beacon-enabled communications, the PAN coordinator periodically broadcasts the beacon frames. However, in the case of non-beacon-enabled communication, the PAN coordinator randomly sends beacon frame. The PAN device has to wait until the channel is available in idle state; otherwise it waits for a random amount of time.

Whenever a device is not sending data, the ZigBee protocol turns off its radio interface to save the energy. Similarly, a node can be in the active state whenever it is transmitting a beacon message. The beacon interval depends on the data rate, and it is usually ranging between 15.36 ms and 251.65 ms at 250 kbits/s. Keeping in view the communication model, we proposed that the BLE is an excellent choice for using it in the room coverage while using ZigBee for the entire house.

**3.5. Proposed Smart Building System.** The proposed system architecture is composed of different embedded devices, such as mobile sensors, healthcare system, and security system, located in a large geographic area. Also, these embedded devices are deployed in home, police station, and fire and brigade centers. These devices are interconnected with each other with the help of Internet using short range communication technologies, that is, Bluetooth and ZigBee, or long range technologies such as WLAN and cellular, as shown in Figure 5.

In Figure 5, different electric appliances are interconnected with the help of Internet. These appliances are also connected with the home server that helps in storing the

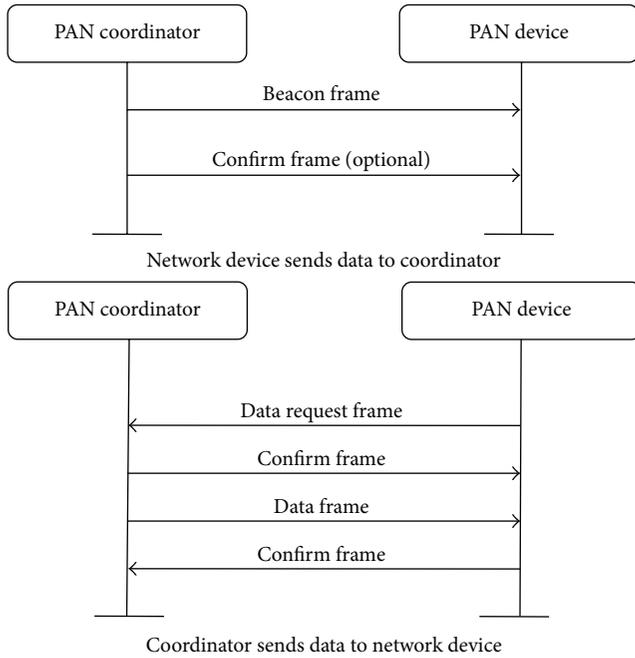


FIGURE 4: Non-beacon-enabled communications in ZigBee protocol.

data for our analysis purposes. Electric appliances include healthcare system for elderly people, wearable devices, social network, and Kinect Xbox for a security system. In the given figure, these devices are connected to the actuator with the help of ZigBee. ZigBee device is then transmitting these signals to the home server. In our proposed system architecture, actuator acts as an intelligent device that provides a medium of communication among various electric appliances using ZigBee or Bluetooth. For instance, various electric appliances, such as smart meter, water usage meter, and lights, are connected to the actuator. These electronic devices perform accordingly when the user is at home or not. Moreover, also, it depicts the level of the water that is used by each user in a home. Similarly, all the electric appliances are controlled remotely. For instance, if a user is not in a home, the user can access home server with the help of 3G/4G data network and can send instructions to the devices to perform a particular task. Moreover, a surveillance system is designed for elderly people and small kids in a home. In the case of elderly people, wearable devices are used that could detect the body gestures of the object and could decide the position of the object using body area network in integration with BLE. Similarly, for small kids in a home, our designed system could be used to find out the current location of the kids. To elaborate the finding of the small kids in a home, the following technique could be applied, which shows the registration phase among coordinator, PMD, and AGW.

As delineated in Figure 6, coordinator is attached to PMD that sends router solicitation (RS) message to PMD. The RS message has the coordinator ID (GDID) and link-layer address (AIDs). The PMD sends a response message to the coordinator upon receiving RS message, and then it generates RA message again back to the coordinator. Afterward, PMD

TABLE I: Home GDID register (HGR).

Number	ID	LOC	Domain
1	GDID1	LLOC (AID) of PMD	Home
2	GDID2	GLOC of AGW	Visiting
3	...	...	...

sends Location Update Request (LUR) to AGW as shown in Table 1.

Upon establishing connection with the AGW, the PMD sends device ID request to the AGW. Afterward, AGW authenticates whether GDID fits in the same domain or not. Since GDID has the information about its home domain, after authentication, location discovery message is sent to the c-AGW. Afterward, c-AGW will first check the HGR mapping table and then will generate response to AGW that contains location discovery message. Upon receiving this message, the AGW adds this information to its mapping table and then sends device ID message to PMD. As a result, data message is forwarded to C-PMD as shown in Figure 7.

The proposed smart building system architecture is equipped with the healthcare embedded system that helps in assisting users in a home. There are various conditions of health that are considered from various datasets, which are diabetics, blood pressure, and other activities, such as climbing stairs; these embedded sensors are attached to the human body that constantly monitor various parameters in a human body, such as diabetes and body temperature. The proposed system architecture welcomes the incoming data that helps in processing the data by the designed algorithms. For example, if a patient condition is dropping or elevating more than a particular threshold, then the system generates alerts and warning messages. These messages and alerts are sent to the home server; the home server is connected to the Internet that sends the data to the concerned department. The sensors attached to the body of the patients communicate locally using the BLE technology. Moreover, the communication with the home server is done using the ZigBee technology. Similarly, the data from the home server is sent to the remote department using WIFI or cellular technology. The concerned department initiates the emergency conditions upon receiving the alert and sends the ambulance or doctor to the patient location. Similarly, other emergency conditions such as in the case of fire, gas leakage, and water pipe leakage can be controlled using the same architecture. The web server plays a vital role in this situation by sending the information to the other web servers for more efficient recovery and timely services.

#### 4. Analytical Results and Discussion

The proposed system architecture is provided with the algorithm that helps in monitoring fire level in the home. Moreover, the designed system is also used for monitoring the patient heartbeat rate in a home. Initially, the system is provided with the information of flashlight and temperature since this information is used for the setup of the initial threshold to check whether the system generates an alarm or

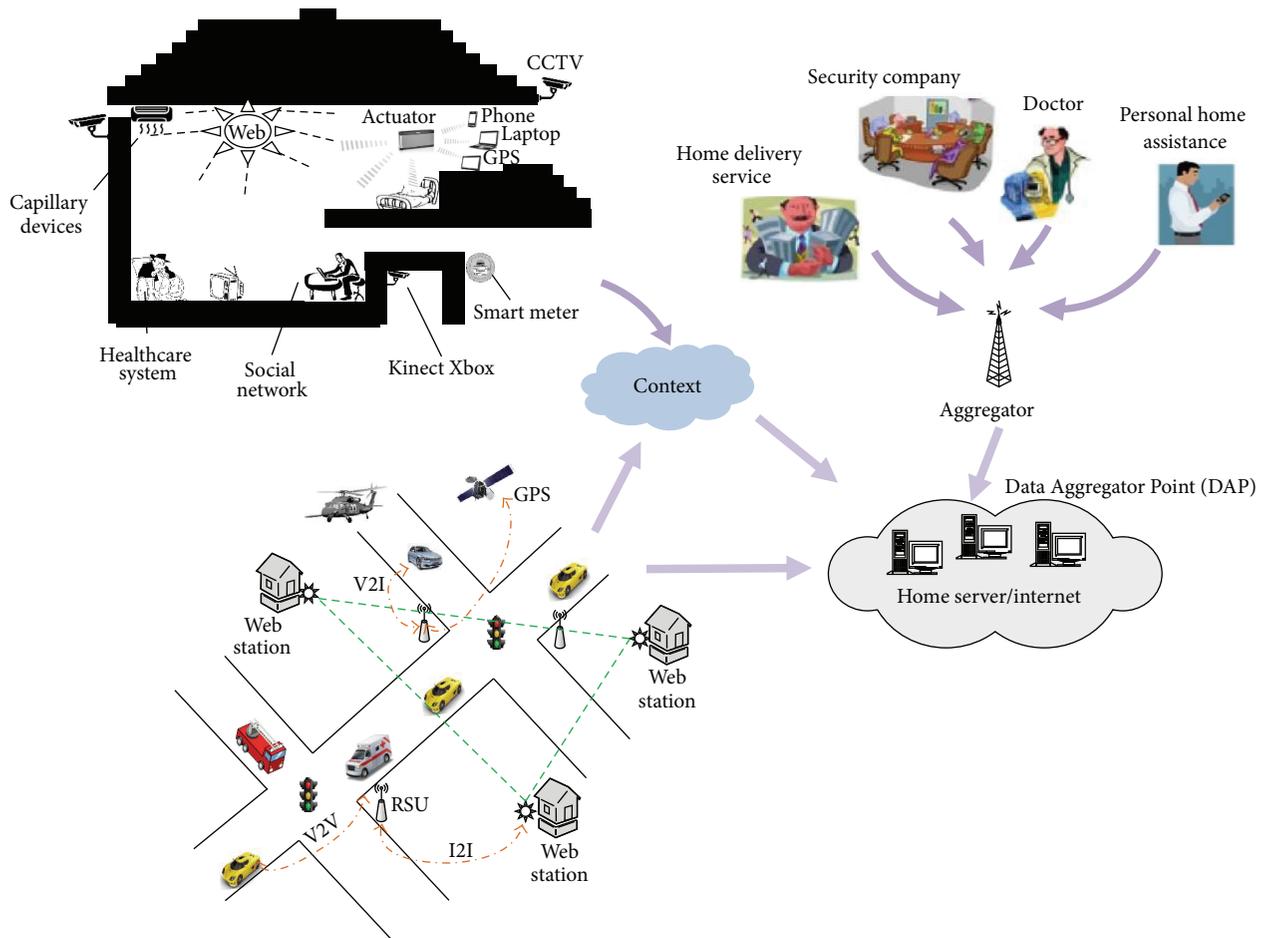


FIGURE 5: Proposed architecture for context-aware low power mobile sensors for sensing discrete events.

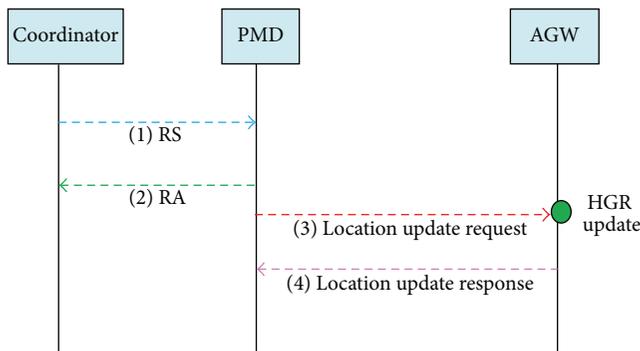


FIGURE 6: Initial registration.

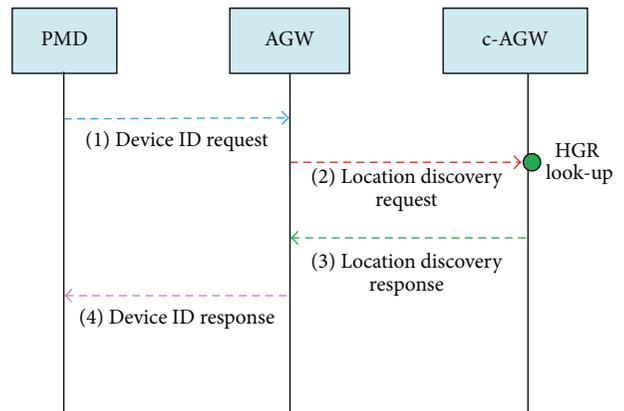


FIGURE 7: Packet delivery operation.

not—depending on the values it received from the embedded sensors. For our analysis, we have tested the proposed algorithm many times in order to check the system accuracy. We have noticed that the proposed system generates zero-degree false alarm. Moreover, if the system is taking measurements and it is within the limits for a longer time, then the system calculates the mean, standard deviation, and maximum value readings. Similarly, the corresponding action is taken while

considering the above values. The same analytical results have been measured in the case of the chronic patient. The thresholds are set on the lower and upper limits of the blood pressure. In case of the higher blood pressure or lower blood pressure as compared with the threshold, the alarm is generated and sent to the concerned department. A pseudocode is given in Algorithm 1.

```

(1) For each (Temperature readings) do
  IF (Temp >  $\delta_{STemp}$  && Flash_light >  $\delta_{FL}$ )
    Fire: Detected;
  Else IF: (Temp >  $\delta_{NTemp}$ )
    Fire = Analyze ( $\bar{x}_{Temp}$ ,  $\sigma^2_{Temp}$ , Max_valTemp)
  Else
    Next();
  Alert();
(2) For each (Heart_Rate) Do
  //Define Thresholds
  Assign  $\delta_{NR}$  and  $\delta_{SR}$  using Table 2
  //Decision
  IF: (Heart_Rate >  $\delta_{SR}$ )
    Alert();
  ELSE IF: (Heart_Rate >  $\delta_{NR}$ )
    Analyze (Heart_Rate)
  ELSE:
    Next();

```

ALGORITHM 1: Fire detection and patient monitoring.

TABLE 2

Age	$\delta_{NR}$	$\delta_{SR}$	Age	$\delta_{NR}$	$\delta_{SR}$
<20	170	200	<50	145	175
<30	162	190	<60	136	170
<40	153	185	>60	128	150

For our analysis, various datasets are considered, such as temperature datasets (118 MB), ECG datasets (227 MB), and heartbeat rate datasets of 1.7 GB. These datasets are taken from [29, 31–33]. We have performed analysis using Hadoop ecosystem by considering MapReduce function. Beside these, other parameters, such as activity parameters and medical health parameters, are taken into account in single datasets. In order to validate the proposed system architecture, we have performed analysis on the above-mentioned datasets to check the throughput and processing time as shown in Figure 8. In this figure, the size of the temperature datasets is taken as minimum; therefore, greater throughput is noticed. However, if we increase the size of the datasets, the throughput of the designed system drastically decreased.

In order to extend our simulation analysis up to further extent, we have considered a scenario of the house with five members. The average energy consumption by the BLE and ZigBee sensors attached to different appliances is calculated for a duration of 12-hour time (duty cycle is 100%). The duty cycle is the time in which the radio interface remains in the active state. The energy efficient communication stack has a duty cycle less than 1%. However, we want to utilize fully the effect of both the technologies; therefore we keep the radio interface always in on state. The sensor embedded in each device operates on a pair of AA batteries (3000 mAh). The BLE-enabled sensor consumes an amount of 6 mA of current in the active state. Similarly, the ZigBee is consuming an amount of 13 mA current. The member of the house randomly moves around the house and randomly switches on/off the

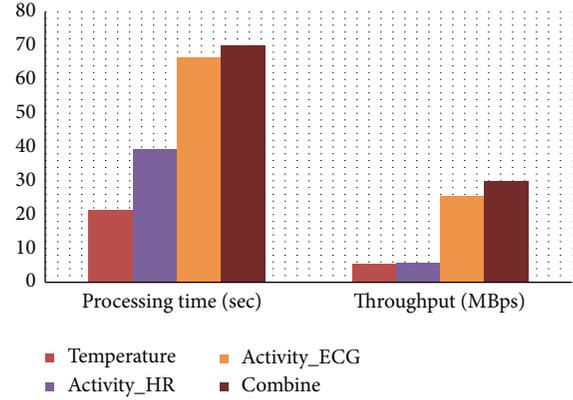


FIGURE 8: Efficiency of proposed system.

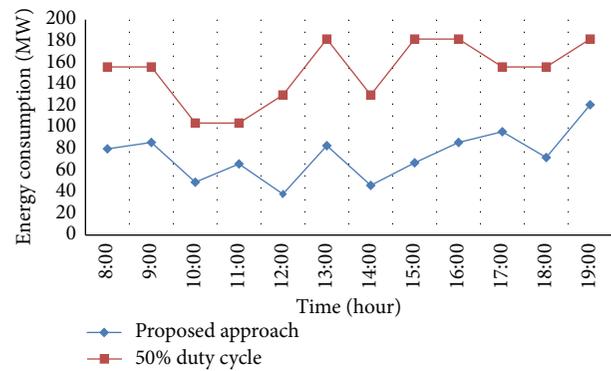


FIGURE 9: Energy consumption while the sensors operate at 50% duty cycle.

device available in the house. The BLE is only used in the room environment while the ZigBee is used in the entire house. The security system and door opening and closing are also controlled through remote control operating via ZigBee. Similarly, different appliances like TV, refrigerator, washing machine, light bulbs, microwave oven, and so forth are also controlled through sensors both using ZigBee and BLE (only rooms). The communication in the range of 50 m is carried out using BLE and more than 100 m using the ZigBee. The average energy consumption values of both technologies for 12-hour time (50% duty cycle) are graphed in Figure 9. Similarly, the total energy consumption of both technologies in the 12-hour time (100% duty cycle) is also computed as shown in Figure 10.

Different types of embedded sensors are deployed in different locations of the home, such as in a kitchen, for human body, and for security services. We have calculated the energy conception of each sensor as shown in Figure 11. The sensors that are deployed in a room usually consume less energy since these sensors are using their transceiver for less duration of time because of the daylight, whereas the usage of these sensors increases in the evening that generally results in increasing the battery usage of the sensors. However, there are some particular times in which the sensor is in the sleep state. At that point, the sensor node does not consume energy.

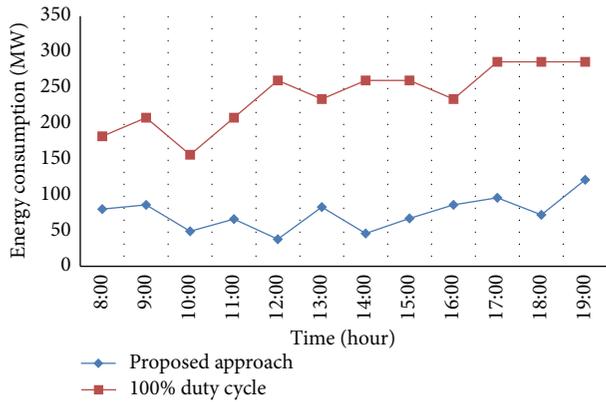


FIGURE 10: Energy consumption while the sensors operate at 100% duty cycle.

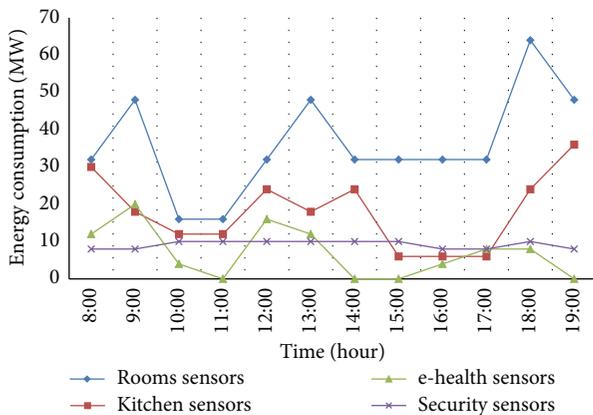


FIGURE 11: Total energy consumption of the sensors during 12-hour time.

Moreover, the duty cycle of sensors is shown in Figure 12. In this figure, the graph shows that less amount of energy is consumed by the sensor because of the daylight. Therefore, the sensor turns off its transceiver. However, as the light becomes dim, the duty cycle of the sensor increases.

## 5. Conclusion

In this paper, we provide a new concept of smart building that is proposed to integrate the concept of the IoT based on the Internet using low power mobile sensors. Furthermore, we define a communication model for sharing data using the same medium. The communication model provides a common medium for the communication of all the heterogeneous devices. Furthermore, a system architecture is also proposed based on the Internet application. The Internet application concept is used to send or receive action message over the network. The proposed architecture provides the application, analysis, and visualization aspects where various devices are integrating various electronic devices. The performance of the system architecture is tested on Hadoop using UBUNTU 14.04 LTS Core™ i5 machine with 3.2 GHz processor and

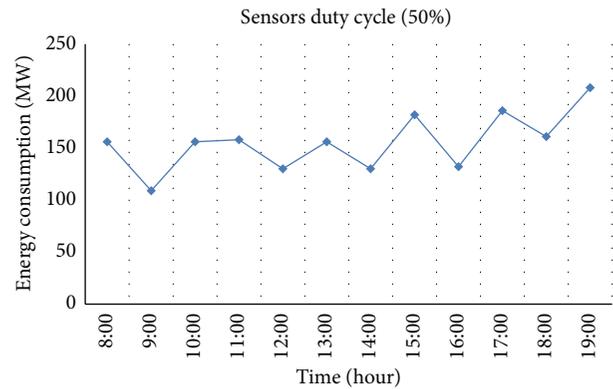


FIGURE 12: Sensors duty cycle.

4 GB memory. Similarly, the energy consumption of the sensors installed in the proposed smart home is also computed. The energy consumption of the sensors using the proposed architecture is significantly reduced. The final evaluations show that the performance of the proposed network architecture fulfills the required desires of the users connected to it, whether the input data is a real-time as well as offline while taking actions at real time. For future work, we are planning to develop a system for disaster management in smart building. The disaster management will be based on the user and appliances' data. The system will decide how to guide its user at the time of disaster.

## Competing Interests

The authors declare that the grant, scholarship, and/or funding mentioned in Acknowledgments does not lead to any conflict of interests. Additionally, the authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This work was supported by the Institute for Information & Communications Technology Promotion (IITP) grant funded by the Korea government (MSIP) (no. 10041145, Self-Organized Software Platform (SoSp) for Welfare Devices). This work is also supported by the National Research Foundation of Korea (NRF) grant funded by the Korean Government (MSIP) (NRF-2013R1A2A1A01014020).

## References

- [1] M. K. Bidhandi, M. K. Bidhandi, and S. A. H. R. Ebrahimi, "Introduce an object-oriented simulator for analyzing discrete events in smart buildings," in *Proceedings of the IEEE International Congress on Technology, Communication and Knowledge (ICTCK '14)*, pp. 1–5, Mashhad, Iran, November 2014.
- [2] T. V. Nguyen, J. G. Kim, and D. Choi, "ISS: the interactive smart home simulator," in *Proceedings of the 11th International Conference on Advanced Communication Technology (ICACT '09)*, pp. 1828–1833, IEEE, Gangwon-Do, Republic of Korea,

- 2009.
- [3] Z. F. Jahromi, A. Rajabzadeh, and A. R. Manashty, "A multi-purpose scenario-based simulator for smart house environments," *International Journal of Computer Science and Information Security*, vol. 9, no. 1, pp. 13–18, 2011.
  - [4] L. Atzori, A. Iera, and G. Morabito, "Making things socialize in the internet—does it help our lives?" in *Proceedings of the ITU Kaleidoscope: The Fully Networked Human?—Innovations for Future Networks and Services*, Cape Town, South Africa, December 2011.
  - [5] I. Farris, A. Iera, and S. C. Spinella, "Introducing a novel 'virtual communication channel' into RFID ecosystems for IoT," *IEEE Communications Letters*, vol. 17, no. 8, pp. 1532–1535, 2013.
  - [6] D. Agarwal and S. K. Prasad, "AzureBench: benchmarking the storage services of the Azure cloud platform," in *Proceedings of the IEEE 26th International Parallel and Distributed Processing Symposium Workshops (IPDPSW '12)*, pp. 1048–1057, IEEE, Shanghai, China, May 2012.
  - [7] S. Dixit and R. Prasad, Eds., *Technologies for Home Networking*, John Wiley & Sons, New York, NY, USA, 2007.
  - [8] X. Li, R. Lu, X. Liang, X. Shen, J. Chen, and X. Lin, "Smart community: an internet of things application," *IEEE Communications Magazine*, vol. 49, no. 11, pp. 68–75, 2011.
  - [9] E. Witrant, S. Mocanu, and O. Sename, "A hybrid model and MIMO control for intelligent buildings temperature regulation over WSN," in *Proceedings of the 8th IFAC Workshop on Time-Delay Systems (TDS '09)*, pp. 420–425, September 2009.
  - [10] E. Asimakopoulou and N. Bessis, "Buildings and crowds: forming smart cities for more effective disaster management," in *Proceedings of the 5th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS '11)*, pp. 229–234, IEEE, Seoul, South Korea, July 2011.
  - [11] Q.-S. Jia, Q. Zhao, H. Darabi et al., "Smart building technology," *IEEE Robotics & Automation Magazine*, vol. 21, no. 2, pp. 18–20, 2014.
  - [12] E.-K. Lee, P. Chu, and R. Gadh, "Fine-grained access to smart building energy resources," *IEEE Internet Computing*, vol. 17, no. 6, pp. 48–56, 2013.
  - [13] A. Ahmad, A. Paul, M. Mazhar Rathore, and H. Chang, "Smart cyber society: integration of capillary devices with high usability based on Cyber-Physical System," *Future Generation Computer Systems*, vol. 56, pp. 493–503, 2016.
  - [14] A. Ahmad, A. Paul, and M. Mazhar Rathore, "An efficient divide-and-conquer approach for big data analytics in machine-to-machine communication," *Neurocomputing*, vol. 174, pp. 439–453, 2016.
  - [15] M. K. Bidhandi, M. K. Bidhandi, and S. A. H. R. Ebrahimi, "Introduce an object-oriented simulator for analyzing discrete events in smart buildings," in *Proceedings of the International Congress on Technology, Communication and Knowledge (ICTCK '14)*, pp. 1–5, IEEE, Mashhad, Iran, November 2014.
  - [16] T. Van Nguyen, J. G. Kim, and D. Choi, "ISS: the interactive smart home simulator," in *Proceedings of the 11th International Conference on Advanced Communication Technology (ICACT '09)*, vol. 3, pp. 1828–1833, February 2009.
  - [17] E. G. Pinho and F. H. de Carvalho Junior, "An object-oriented parallel programming language for distributed-memory parallel computing platforms," *Science of Computer Programming*, vol. 80, pp. 65–90, 2013.
  - [18] Q. Fu, P. Li, C. Chen, L. Qi, Y. Lu, and C. Yu, "A configurable context-aware simulator for smart home systems," in *Proceedings of the 6th International Conference on Pervasive Computing and Applications (ICPCA '11)*, pp. 39–44, Port Elizabeth, South Africa, October 2011.
  - [19] A. Rajabzadeh, A. R. Manashty, and Z. F. Jahromi, "A generic model for smart house remote control systems with software and hardware simulators," in *Proceedings of the 5th Conference on Information and Knowledge Technology (IKT '13)*, pp. 262–267, Shiraz, Iran, May 2013.
  - [20] M. Drăgoicea, L. Bucur, and M. Pătrașcu, "A service oriented simulation architecture for intelligent building management," in *Exploring Services Science*, J. Falcão e Cunha, M. Snene, and H. NÓvoa, Eds., vol. 143 of *Lecture Notes in Business Information Processing*, pp. 14–28, Springer, Berlin, Germany, 2013.
  - [21] M. Yoon, S. Chen, and Z. Zhang, "Minimizing the maximum firewall rule set in a network with multiple firewalls," *IEEE Transactions on Computers*, vol. 59, no. 2, pp. 218–230, 2010.
  - [22] O. Rottenstreich, R. Cohen, D. Raz, and I. Keslassy, "Exact worst case TCAM rule expansion," *IEEE Transactions on Computers*, vol. 62, no. 6, pp. 1127–1140, 2013.
  - [23] 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.
  - [24] C. L. Forgy, *On the efficient implementation of production systems [Ph.D. thesis]*, Department of Computer Science, Carnegie Mellon University, Pittsburgh, Pa, USA, 1979.
  - [25] D. P. Miranker, "TREAT: a better match algorithm for AI production system matching," in *Proceedings of the 6th National Conference on Artificial Intelligence (AAAI '87)*, pp. 10–18, 1987.
  - [26] A. Paul, A. Daniel, A. Ahmad, and S. Rho, "Cooperative cognitive intelligence for internet of vehicles," *IEEE Systems Journal*, 2015.
  - [27] A. Paul, "Real-time power management for embedded M2M using intelligent learning methods," *ACM Transactions on Embedded Computing Systems*, vol. 13, no. 5, article 148, 2014.
  - [28] E. N. Hanson and M. S. Hasan, "Gator: an optimized discrimination network for active database rule condition testing," Tech. Rep. TR-93-036, CIS Department, University of Florida, Gainesville, Fla, USA, 1993.
  - [29] A. Reiss and D. Stricker, "Creating and benchmarking a new dataset for physical activity monitoring," in *Proceedings of the 5th International Conference on Pervasive Technologies Related to Assistive Environments*, article 40, ACM, 2012.
  - [30] J. Kim, J. Lee, H. K. Kang, D. S. Lim, C. S. Hong, and S. Lee, "An ID/Locator separation-based mobility management architecture for WSNs," *IEEE Transactions on Mobile Computing*, vol. 13, no. 10, pp. 2240–2254, 2014.
  - [31] M. Zhou and Z.-L. Nie, "Analysis and design of ZigBee MAC layers protocol," in *Proceedings of the International Conference on Future Information Technology and Management Engineering (FITME '10)*, pp. 211–215, Changzhou, China, October 2010.
  - [32] A. Reiss and D. Stricker, "Introducing a new benchmarked dataset for activity monitoring," in *Proceedings of the 16th International Symposium on Wearable Computers (ISWC '12)*, pp. 108–109, IEEE, Newcastle, UK, June 2012.
  - [33] O. Banos, R. Garcia, J. A. Holgado-Terriza et al., "mHealthDroid: a novel framework for agile development of mobile health applications," in *Ambient Assisted Living and Daily Activities*, pp. 91–98, Springer, Berlin, Germany, 2014.

## Research Article

# Noise Estimation and Suppression Using Nonlinear Function with *A Priori* Speech Absence Probability in Speech Enhancement

Soojeong Lee<sup>1</sup> and Gangseong Lee<sup>2</sup>

<sup>1</sup>School of Electronic Engineering, Hanyang University, 222 Wangsimni-ro, Seongdong, Seoul 133-791, Republic of Korea

<sup>2</sup>Kwangwoon University, 20 Kwangwoon-ro, Nowon-gu, Seoul, Republic of Korea

Correspondence should be addressed to Soojeong Lee; leesoo86@hanyang.ac.kr

Received 7 December 2015; Accepted 6 April 2016

Academic Editor: Marco Anisetti

Copyright © 2016 S. Lee and G. Lee. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes a noise-biased compensation of minimum statistics (MS) method using a nonlinear function and *a priori* speech absence probability (SAP) for speech enhancement in highly nonstationary noisy environments. The MS method is a well-known technique for noise power estimation in nonstationary noisy environments; however, it tends to bias noise estimation below that of the true noise level. The proposed method is combined with an adaptive parameter based on a sigmoid function and *a priori* SAP for residual noise reduction. Additionally, our method uses an autoperparameter to control the trade-off between speech distortion and residual noise. We evaluate the estimation of noise power in highly nonstationary and varying noise environments. The improvement can be confirmed in terms of signal-to-noise ratio (SNR) and the Itakura-Saito Distortion Measure (ISDM).

## 1. Introduction

Noise estimation algorithms are essential components of many modern mobile communication, speech recognition, and human computer interaction systems for speech enhancement [1, 2]. It is generally included as a part of the speech enhancement to improve the speech intelligibility or quality of a signal corrupted by noise. However, it is difficult to reduce noise without distorting speech because the performance of any noise estimation algorithm usually depends on a trade-off between speech distortion and noise reduction.

Current single microphone speech enhancement methods belong to two groups, namely, time domain methods such as the subspace method and frequency domain methods such as the spectral subtraction (SS) [3] and minimum mean square error (MMSE) estimator [4]. Both methods have their own advantages and drawbacks. Subspace methods provide a mechanism to control the trade-off between speech distortion and residual noise, but with the cost of a heavy computational load [5]. Frequency domain methods, on the other hand, usually consume less computational resources

but do not have a theoretically established mechanism to control trade-off between speech distortion and residual noise. Among them, spectral subtraction (SS) is computationally efficient and has a simple mechanism to control trade-off between speech distortion and residual noise but suffers from a notorious artifact known as musical noise [6]. These spectral noise reduction algorithms require an estimate of the noise spectrum, which can be obtained from speech absence frames indicated by a voice activity detector (VAD) or, alternatively, with the minimum statistic (MS) methods [7], that is, by tracking spectral minima in each frequency band.

Several recent studies have proposed noise estimation schemes for unknown noise signals [1–14]. The minimum statistics (MS) noise estimation scheme [7] is one that works well in nonstationary noisy environments. Martin proposed an algorithm for noise estimation based on minimum statistics [7]. The ability to track varying noise levels is a prominent feature of the minimum statistics (MS) algorithm [7]. The noise estimate is obtained as the minima values of a smoothed power estimate of the noisy signal, multiplied by a factor that compensates the bias. However, the MS algorithm still has

a tendency to bias the noise estimate below that of the true noise level, regardless of the number of frames [8]. Therefore, it leaves residual noise in the frames of speech absence and in the frames of variation of noise characteristic in highly nonstationary noisy environments.

To solve this problem, we propose a combined adaptive factor based on a sigmoid function and *a priori* speech absence probability (SAP) estimation [9] for biased compensation. Specifically, we apply the adaptive factor  $\delta$  as *a posteriori* SNR. When the *a posteriori* SNR decreases,  $\delta$  increases but is constrained to take a value between  $\delta_{\min}$  and  $\delta_{\max}$ . Thus, the proposed adaptive biased compensation factor  $\delta$  approaches  $\delta_{\max}$  at times when the SNR is low. In addition, when the *a priori* SAP equals unity, the adaptive biased compensation factor  $\delta$  also approaches  $\delta_{\max}$  in each frequency bin and vice versa. Furthermore, our method uses another adaptive parameter to control the trade-off between speech distortion and residual noise for suppressing the estimated noise in highly nonstationary and various noisy environments. The autocontrol parameter is controlled by *a posteriori* signal-to-noise ratio (SNR) as the variation of the noise level.

We evaluate the performance of the proposed algorithm for nonstationary noise and various noise environments. The improvement can be confirmed in the segmental SNR and the Itakura-Saito Distortion Measure (ISDM) [15]. The results show that our proposed method is superior to the conventional MS approach. The structure of the paper is as follows. Section 2 reviews the minimum statistics and the *a priori* SAP estimation algorithms. Section 3 addresses noise estimation and suppression using a linear and a nonlinear function. In Section 4, we express the combined sigmoid function using the *a posteriori* SNR and *a priori* SAP estimation for robust biased compensation. In Section 5, we discuss the experimental results.

## 2. Minimum Statistics (MS) and Speech Absence Probability (SAP)

**2.1. Review of MS.** The noisy speech signal  $y(n)$  can be represented as  $y(n) = x(n) + d(n)$ , where  $x(n)$  is the clean speech signal and  $d(n)$  is the noise signal. Dividing the signal into overlapping frames using a window function and applying the short-time Fourier transform (STFT) [16] to each frame yield the time-frequency representation  $Y(k, l) = X(k, l) + D(k, l)$ , where  $k = 1, 2, \dots, K$  is the frequency bin index and  $l = 1, 2, \dots, L$  is the time frame index. It can be shown that

$$|Y_k(l)|^2 \approx |X_k(l)|^2 + |D_k(l)|^2, \quad (1)$$

where  $|Y_k(l)|^2$ ,  $|X_k(l)|^2$ , and  $|D_k(l)|^2$  are the power spectrum of the noisy speech signal, clean speech, and noise, respectively.

The MS algorithm relies on the fact that the noisy power spectrum often becomes equal to the noise power spectrum during periods of speech pauses [7, 13, 17]. Therefore, an estimate of the noise power spectrum is obtained by separately tracking the minimum of the noisy speech in each frequency

bin. In addition, because the minimum is biased towards lower values, an unbiased estimate may be obtained through multiplication by a bias factor, which is derived from the statistics of the local minimum. To search for the minimum, we take the first-order recursive of the noisy power spectrum:

$$S_k(l) = \alpha S_k(l-1) + (1-\alpha) |Y_k(l)|^2, \quad (2)$$

where  $S_k(l)$  is the smoothed periodogram and  $\alpha$  is the smoothing factor. The smoothing factor used in (2) must be close to 1 to keep the variance of the minimum tracking as small as possible. Hence, time and frequency dependence are required to determine if speech is present or absent. The smoothing factor is therefore derived by minimizing the mean square error between  $S_k(l)$  and  $\sigma_{d,k}^2(l)$ :

$$E \left\{ \left( S_k(l) - \sigma_{d,k}^2(l) \right)^2 \mid S_k(l-1) \right\}, \quad (3)$$

where  $\sigma_{d,k}^2(l)$  is the noise variance:

$$S_k(l) = \alpha_k(l) S_k(l-1) + (1-\alpha_k(l)) |Y_k(l)|^2. \quad (4)$$

In (4), the time-frequency dependent smoothing factor  $\alpha_k(l)$  is used instead of the fixed  $\alpha$  defined in (2). Substituting (4) into (3) and setting the first derivative to 0, we find the optimum value for  $\alpha_k(l)$

$$\alpha_{\text{opt},k}(l) = \frac{1}{1 + \left( S_k(l-1) / \sigma_{d,k}^2(l) - 1 \right)}. \quad (5)$$

According to (5), the smoothing factor can vary between 0 and 1, but such a smoothing factor is not practical [15]. The value of  $\alpha_{\text{opt}}$  becomes progressively smaller for a large *a posteriori* SNR  $\bar{\gamma} \approx (S_k(l-1) / \sigma_{d,k}^2(l))$  (speech present). However, smoothing is required even during periods of speech because the speech power spectrum also contains a percentage of noise. Hence, the smoothing factor has a floor of (0.3), which results in a maximum of only (70%) of the original spectrum remaining within any one frame. Conversely, when the *a posteriori* SNR  $\bar{\gamma}$  is low (speech is absent)  $\alpha$  tends towards 1, which causes the smoothed output to lock onto the previous value. To eliminate this, (5) is multiplied by  $\alpha_{\text{max}} = 0.96$ . From (5), we note that  $\alpha_{\text{opt},k}(l)$  depends on the true noise variance  $\sigma_{d,k}^2(l)$ , which is unknown. In practice, we can replace  $\sigma_{d,k}^2(l)$  with the latest estimated value  $\hat{\sigma}_{d,k}^2(l-1)$ . In general, however, this lags the true noise variance, and hence the estimated smoothing factor may be too small or large. Problems may arise when  $\alpha_{\text{opt},k}(l)$  is close to 1 because  $S_k(l)$  will not respond fast enough to changes in the noise. Thus, tracking errors were monitored in [7] by comparing the average short-term smoothed periodogram to the estimated noise variance. After including the correction factor [7]

$$\alpha_c(l) = \frac{1}{1 + \left( \sum_{k=1}^K S_k(l-1) / \sum_{k=1}^K |Y_k(l)|^2 - 1 \right)^2}, \quad (6)$$

the final factor

$$\alpha_{\text{opt},k}(l) = \frac{\alpha_{\text{max}} \cdot \alpha_c(l)}{1 + (S_k(l-1)/\hat{\sigma}_{d,k}^2(l-1) - 1)^2} \quad (7)$$

is also smoothed over time [7].

The estimated noise power based the MS algorithm [7] is obtained by searching for a minimum within a finite window length  $C$  of the smoothed power estimates  $P(k, l)$ :

$$S_{\text{min},k}(l) = \min \{S_k(l), S_k(l-1), \dots, S_k(l-C)\}. \quad (8)$$

Because the minimum power estimate obtained through the time-varying smoothing factor is smaller than the mean value, the MS algorithm requires a bias compensation for the unbiased noise power estimate as detailed in the following [7]:

$$\hat{\sigma}_{d,k}^2(l) = \beta_{\text{min},k}(l) \cdot S_{\text{min},k}(l), \quad (9)$$

where  $\hat{\sigma}_{d,k}^2(l)$  is the unbiased noise power estimate. The quantity  $\beta_{\text{min},k}(l)$  is the bias compensation factor.

**2.2. Review of Speech Absence Probability.** The two-state model of speech events can be represented as a binary hypothesis model [9, 15, 17]:

$$\begin{aligned} H_0(k, l) : Y(k, l) &= D(k, l), \\ H_1(k, l) : Y(k, l) &= X(k, l) + D(k, l), \end{aligned} \quad (10)$$

where  $H_0(k, l)$  and  $H_1(k, l)$  represent the absence and presence of speech, in the  $k$ th frequency bin of the  $l$ th frame, respectively, and where

$$P(H_0(k, l) \equiv q(k, l)) \quad (11)$$

is the *a priori* probability that speech will be absent. An efficient estimator is derived for the *a priori* SAP using a soft-decision approach based on the estimated *a priori* SNR [9]. A recursive average of this can be defined as

$$\zeta(k, l) = \beta \zeta(k, l-1) + (1 - \beta) \hat{\xi}(k, l-1), \quad (12)$$

where  $\beta$  is a time constant. The decision-directed method proposed by Ephraim and Malah [4] provides a useful estimation scheme for the *a priori* SNR:

$$\hat{\xi}(k, l) = a \frac{\hat{X}^2(k, l-1)}{\hat{\sigma}_d^2(k, l)} + (1 - a) \max[\gamma(k, l) - 1, 0], \quad (13)$$

where  $a$  ( $0 < a < 1$ ) is a smoothing factor,  $\max$  is a function that prevents negative values, and  $\gamma(k, l) \approx |Y(k, l)|^2/\hat{\sigma}_d^2(k, l)$  represents the *a posteriori* SNR [9]. The local and global averaging window are then applied to (13) [9], resulting in

$$\zeta_\lambda(k, l) = \sum_{i=-w_\lambda}^{w_\lambda} h_\lambda(i) \zeta(k-i, l), \quad (14)$$

where the subscript  $\lambda$  may denote either “local” or “global” window and  $h_\lambda$  is a normalized window of size  $2w_\lambda + 1$ . We

define two parameters  $P_{\text{local}}$  and  $P_{\text{global}}$ , which represent the relationship between the above averages and the likelihood of speech in the  $k$ th frequency bin of the  $l$ th frame. These parameters are given as [9]

$$P_\lambda(k, l) = \begin{cases} 0, & \text{if } \zeta_\lambda(k, l) \leq \zeta_{\text{min}}, \\ 1, & \text{if } \zeta_\lambda(k, l) \geq \zeta_{\text{max}}, \\ \frac{\log(\zeta_\lambda(k, l)/\zeta_{\text{min}})}{\log(\zeta_{\text{max}}/\zeta_{\text{min}})}, & \text{otherwise,} \end{cases} \quad (15)$$

where  $\zeta_{\text{min}}$  and  $\zeta_{\text{max}}$  are empirical constants, maximized to attenuate noise while leaving weak speech components unaffected. The third parameter  $P_{\text{frame}}(l)$ , which is required to attenuate more noise in speech-absent frames, is based on the speech energy in neighboring frames [9]:

$$\begin{aligned} &\text{If } \zeta_{\text{frame}}(l) > \zeta_{\text{min}} \text{ then} \\ &\quad \text{if } \zeta_{\text{frame}}(l) > \zeta_{\text{frame}}(l-1) \text{ then} \\ &\quad \quad P_{\text{frame}}(l) = 1 \\ &\quad \quad \zeta_{\text{peak}}(l) = \min\{\max[\zeta_{\text{frame}}, \zeta_{p\text{min}}], \zeta_{p\text{max}}\} \\ &\quad \text{else} \\ &\quad \quad P_{\text{frame}}(l) = \mu(l) \end{aligned}$$

**Else**

$$P_{\text{frame}}(l) = 0,$$

where  $\zeta_{\text{frame}}(l) = (1/(K/2 + 1)) \sum_{k=1}^{K/2+1} \zeta(k, l)$  is an average in the frequency domain,  $\mu(l)$  represents a soft transition from speech to noise,  $\zeta_{\text{peak}}$  is a confined peak value of  $\zeta_{\text{frame}}$ , and  $\zeta_{p\text{min}}$  and  $\zeta_{p\text{max}}$  are empirical constants that determine the delay of the transition, as defined in [9]. Finally, the *a priori* SAP can be defined as [9]

$$\hat{q}(k, l) = 1 - P_{\text{local}}(k, l) \cdot P_{\text{global}}(k, l) \cdot P_{\text{frame}}(l). \quad (16)$$

Accordingly,  $\hat{q}(k, l)$  is larger if either previous frames or recent neighboring frequency bins do not contain speech. Therefore, when SAP goes to 1, the speech presence probability goes to 0.

### 3. Noise Estimation and Suppression Using Linear and Nonlinear Function

**3.1. Combining Adaptive Factor Based on Sigmoid Function and A Priori SAP.** In this section, we propose a method that combines the adaptive factor based on the sigmoid function and the *a priori* SAP estimation [9] to achieve biased compensation.

First, we can detect the adaptive factor by requiring the smoothed power spectrum  $P(k, l)$  be equal to the updated noise power estimator  $\hat{\sigma}_d^2$  during speech absence region. In particular, we can determine the adaptive factor by minimizing the mean squared error (MSE) between  $P(k, l)$  and  $\hat{\sigma}_d^2$  as follows:

$$E \left\{ \left( P(k, l) - \hat{\sigma}_d^2(k, l) \right)^2 \mid P(k, l-1) \right\}, \quad (17)$$

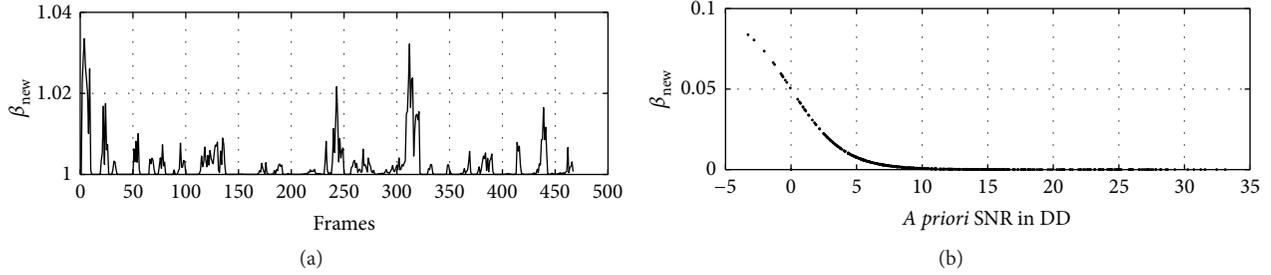


FIGURE 1: (a) Plot of the adaptive factor  $\delta$  in the frame index. (b) Adaptive factor  $\delta$  using a sigmoid function based on the *a posteriori* SNR.

where we assume that the updated noise power estimator  $\bar{\sigma}_d^2$  during the speech absence region is

$$\bar{\sigma}_d^2(k, l)^2 \approx \hat{\sigma}_d^2(k, l)^2 + \delta(k, l). \quad (18)$$

Substituting (18) into (17) then after taking the first derivative of the MSE with respect to  $\delta(l)$  and setting it equal to zero, we get the adaptive factor for  $\delta(l)$ :

$$\delta(l) = P(k, l) - \hat{\sigma}_d^2(k, l), \quad (19)$$

where  $\hat{\sigma}_d^2$  is the unbiased noise power estimate in (9). We apply the adaptive factor based on the sigmoid function to the biased compensation factor of the MS algorithm according to the *a posteriori* SNR:

$$\delta(l) \approx \eta \cdot \frac{1}{(1 + \exp(-(-\rho \cdot \text{SNR}(l))))}, \quad (20)$$

where  $\delta(l)$  is derived from the slope factor  $\rho = 0.5$  and the empirical constant  $\eta = 0.1$  for  $\delta_{\max}$ . The *a posteriori* SNR is

$$\text{SNR}(l) = 10 \cdot \log \left( \frac{\| |Y(k, l)|^2 \|}{\| \hat{\sigma}_d^2(k, l) \|} \right), \quad (21)$$

where  $\| \cdot \|$  is the Euclidean length of a vector. The adaptive factor  $\delta(l)$  is controlled by the *a posteriori* SNR. When the *a posteriori* SNR decreases,  $\delta(l)$  increases but is constrained to take a value between  $\delta_{\min}$  and  $\delta_{\max}$ . Thus, the proposed adaptive biased compensation factor  $\delta(l)$  approaches  $\delta_{\max}$  at times when the SNR is low. In addition, when the *a priori* SAP equals unity, the adaptive biased compensation factor  $\delta(l)$  is also equal to  $\delta_{\max}$  in each frequency bin and vice versa. The adaptive factor is shown to be a biased compensation in Figure 1. It shows, as suggested by (20) and (21), that as the *a posteriori* SNR increases,  $\delta(l)$  decreases but  $\delta(l)$  maintains a value between  $\delta_{\max}$  ( $\delta_{\max} \ll 0.1$ ) and  $\delta_{\min}$ . Thus, the adaptive factor  $\delta(l)$  approaches  $\delta_{\min}$  when the SNR is close to 20 dB. Simulation results show that an increase in the  $\delta(l)$  is good for noisy signals with a low SNR of less than 5 dB and that a decrease in  $\delta(l)$  is good for noisy signals with a relatively high SNR greater than 10 dB. We can thus control the trade-off between speech distortion and residual noise in the frame index using  $\delta(l)$ . In (22), let  $\bar{\sigma}_d^2(k, l)$  be the updated noise power estimate according to the combined *a priori* SAP and the adaptive factor:

$$\bar{\sigma}_d^2(k, l) = \hat{\sigma}_d^2(k, l) + \delta(l) \cdot \hat{q}(k, l). \quad (22)$$

The term  $\hat{q}(k, l)$  is the *a priori* SAP in (16). When  $\hat{q}(k, l)$  becomes 1, the adaptive biased compensation factor  $\delta(l)$  is equal to  $\delta_{\max}$ . Therefore, the speech absence region is efficiently compensated by combining the *a priori* SAP and the adaptive factor in the  $k$ th frequency bin of the  $l$ th frame. As a result, the updated noise power estimator for the optimal smoothing factor  $\hat{\alpha}_{\text{opt}}(k, l)$  of  $\alpha_{\text{opt}}(k, l)$  is deduced from (7) as

$$\hat{\alpha}_{\text{opt}}(k, l) = \frac{\alpha_{\max} \cdot \alpha_c(l)}{1 + (P(k, l-1) / \bar{\sigma}_d^2(k, l-1) - 1)^2}. \quad (23)$$

**3.2. Estimated Noise Suppression Using Linear Function.** In this subsection, our method uses another adaptive parameter to control the trade-off between speech distortion and residual noise for suppressing the estimated noise in a highly nonstationary and varying noisy environment. The autocontrol parameter is controlled by *a posteriori* signal-to-noise ratio (SNR) as the variation of the noise level.

The estimated clean speech power spectrum can be represented as shown in (28). One has

$$\text{SNR}(k) = \log \left( \frac{P(k, l)}{\bar{\sigma}_d^2(k, l)} \right), \quad (24)$$

$$\zeta_s = \frac{\zeta_{\min} - \zeta_{\max}}{\text{SNR}_{\max} - \text{SNR}_{\min}}, \quad (25)$$

$$\zeta_o = \zeta_{\max} - \zeta_s \cdot \text{SNR}_{\max}, \quad (26)$$

$$\zeta(k) = \zeta_s \cdot \text{SNR}(k) + \zeta_o, \quad (27)$$

$$|\hat{X}(k, l)|^2 \approx |Y(k, l)|^2 - \zeta(k) \cdot \bar{\sigma}_d^2(k, l), \quad (28)$$

where  $\zeta(k)$  is the oversubtraction factor,  $\zeta_s$  is the slope, and  $\zeta_o$  is the offset. The constants  $\zeta_{\min} = 1$ ,  $\zeta_{\max} = 3$ ,  $\text{SNR}_{\max} = 20$  dB, and  $\text{SNR}_{\min} = -5$  dB, respectively [3]. The adaptive linear factor  $\zeta(k)$  affects the amount of speech distortion caused by the spectral subtraction in (28). The factor  $\zeta(k)$  offers a large amount of flexibility to the modified spectral subtraction (MSS) scheme. The  $\text{SNR}(k)$  in (24) is the *a posteriori* SNR in frequency bin. The estimated clean speech signal can then be transformed back to the time domain by taking the inverse STFT and synthesizing using the overlap-add method.

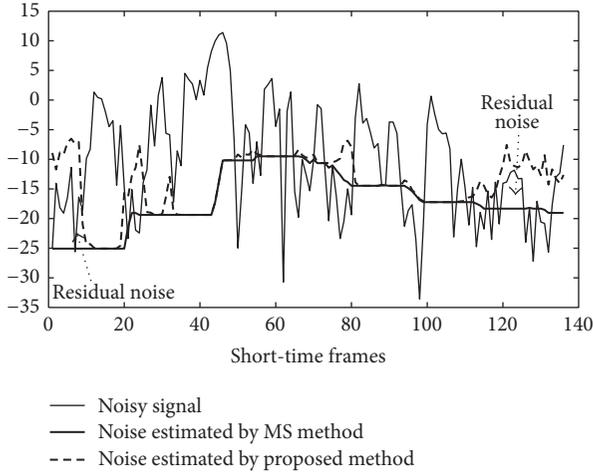


FIGURE 2: Comparison between the noisy signal, noise estimated by MS, and noise estimated by the proposed method in restaurant 5 dB noise environment.

#### 4. Experimental Results and Discussion

The noisy signals used in our evaluation were taken from the NOIZEUS database [15]. We used 30 test utterances, of which three each were from male and female speech signals. The analyzed signal was sampled at 8 kHz and short-time Fourier-transformed using 50% overlapping Hamming windows of 256 samples. Both the MS [7] and proposed methods track the minimum of the noisy speech to update the noise estimate in Figure 2. The MS method is obtained by tracking the minimum of the noisy power spectrum over a specified number of frames. Thus, the MS algorithm noise estimate tends to be biased below the true noise level, regardless of the number of frames. Our proposed method efficiently compensates the speech absence region by combining the adaptive bias compensation factor and *a priori* SAP. This implies that the proposed method is more accurate than the conventional one and could improve residual noise reduction.

Figure 3 shows the clear superiority of the proposed method in highly nonstationary noisy environments. The conventional method [7] does not work well from initial frame to 20 frames of car noise (15 dB) and from 110 frames to 130 frames of car (15 dB) and also suffered from residual noise. A different outcome is observed in the red circle of Figure 3. Particularly, the robust characteristics of the proposed method in spite of the variation of the noisy environments are well demonstrated. Thus, we can estimate more exactly the noise level to reduce a residual noise when compared with conventional method in highly nonstationary noisy environments.

The spectrum of the clean signal is given in Figure 4(a), and the spectrum of the noisy speech signal for speech enhancement using the MS plus spectral subtraction (SS) (MS + SS) [3, 7] method is given in Figure 4(b). We can also observe the minimum controlled recursive averaging (MCRA) with SS in Figure 4(c). There is residual noise in Figure 4(c) from  $0\text{ s} < t < 0.15\text{ s}$  and at  $t > 1.8\text{ s}$ , partly

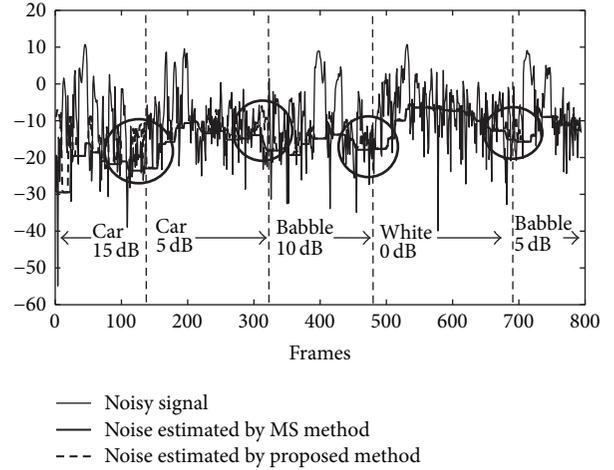


FIGURE 3: Comparison between the noisy signal, noise estimated by minimum statistics (MS), and noise estimated by the proposed method in highly nonstationary noisy environments.

TABLE 1: Objective evaluation and comparison of the proposed method segmental SNR values.

Noise	Method	SNR			
		0 (dB)	5 (dB)	10 (dB)	15 (dB)
White	MS	4.27	8.77	12.83	16.57
	MCRA	5.08	9.99	13.56	17.15
	Proposed	<b>5.69</b>	<b>10.64</b>	<b>14.09</b>	<b>17.40</b>
Car	MS	3.44	7.48	12.01	16.10
	MCRA	4.92	7.93	11.85	16.42
	Proposed	<b>5.39</b>	<b>8.84</b>	<b>12.45</b>	<b>17.06</b>
Babble	MS	1.83	6.00	11.17	15.03
	MCRA	3.73	6.79	10.52	16.22
	Proposed	<b>3.83</b>	<b>7.05</b>	<b>11.60</b>	<b>16.23</b>
Airport	MS	1.75	7.04	9.85	14.73
	MCRA	1.64	7.54	9.66	15.15
	Proposed	<b>2.17</b>	<b>8.16</b>	<b>10.10</b>	<b>15.73</b>
Street	MS	2.77	6.75	<b>10.88</b>	<b>15.31</b>
	MCRA	2.34	7.40	9.88	14.36
	Proposed	<b>3.18</b>	<b>8.24</b>	10.62	15.03
Restaurant	MS	0.31	4.48	9.40	14.74
	MCRA	0.27	5.47	9.20	15.24
	Proposed	<b>0.28</b>	<b>5.57</b>	<b>9.43</b>	<b>15.58</b>

because of the inability of the noise estimation algorithm to bias below the true noise level. The spectrogram of the proposed methods for noise reduction is shown in Figure 4(d). In contrast, panel Figure 4(d) shows that the residual noise is more clearly reduced than the conventional methods.

Tables 1 and 2 summarize the averaged results of the segmental SNR and the Itakura-Saito Distortion Measure (ISDM) [15]. The segmental SNR can be evaluated in either the time or frequency domain. The time domain measure is perhaps one of the simplest objective measures used to evaluate speech enhancement method. For this measure to

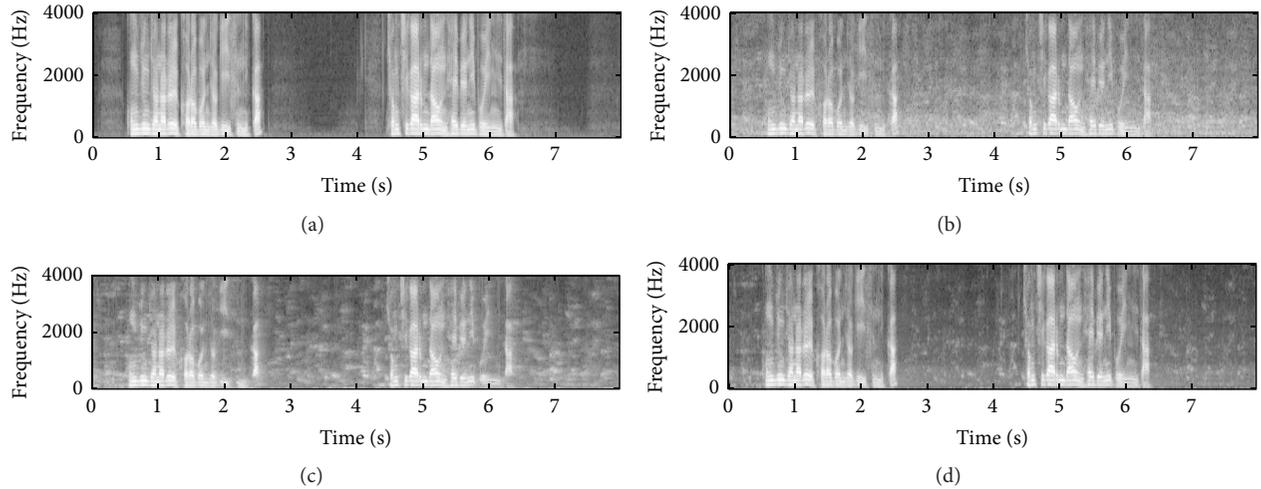


FIGURE 4: Frequency domain results of speech enhancement for exhibition noise 5 dB SNRs in noisy environments. (a) Original spectrogram, (b) spectrogram using MS with SS method, (c) spectrogram using the MCRA with SS method, and (d) spectrogram using the proposed method.

TABLE 2: Objective evaluation and comparison of the Itakura-Saito Distortion Measure (ISDM).

Noise	Method	ISDM			
		0 (dB)	5 (dB)	10 (dB)	15 (dB)
White	MS	1.20	0.87	0.60	0.42
	MCRA	0.92	0.44	0.55	0.37
	Proposed	<b>0.84</b>	<b>0.43</b>	<b>0.38</b>	<b>0.34</b>
Car	MS	0.15	0.16	0.02	0.01
	MCRA	0.21	0.05	0.02	0.01
	Proposed	<b>0.09</b>	<b>0.02</b>	0.02	0.02
Babble	MS	0.15	0.06	0.02	0.01
	MCRA	0.12	0.02	0.01	0.01
	Proposed	<b>0.08</b>	0.02	0.01	0.01
Airport	MS	0.19	0.06	0.02	0.02
	MCRA	0.14	0.04	0.02	0.01
	Proposed	<b>0.10</b>	0.04	<b>0.01</b>	0.01
Street	MS	0.16	0.18	<b>0.06</b>	<b>0.03</b>
	MCRA	0.55	0.13	0.09	0.05
	Proposed	0.17	<b>0.09</b>	0.08	0.04
Restaurant	MS	0.10	0.05	0.01	0.01
	MCRA	0.10	0.03	0.01	0.01
	Proposed	0.10	<b>0.01</b>	0.01	0.01

be meaningful it is important that the original and processed signals be aligned in time and that any phase error present be corrected [15]. For various noise types with an input SNR ranging from 0 to 15 dB, the segmental SNR after processing was clearly better for the proposed method compared to conventional ones [7], except for the case of (highlighted in bold). We can also confirm that our methods work well to control the trade-off between speech distortion and residual noise for suppressing the estimated noise in highly nonstationary and various noisy environments.

The ISDM was shown to give a good correlation with subjective intelligibility measures specifically the diagnostic acceptability measure (DAM). This results in an objective test that can be used to produce a good meaningful result. This also results in a test that shows the distortion and noise reduction [15]. Here, we can confirm that the results of the ISDM with the proposed method produce good results of ISDM when compared with the conventional methods except for the case of the the MS method with SS in street 10 dB noisy signal.

## 5. Conclusion

We presented a modified noise estimation and suppression algorithm that combined the nonlinear function and *a priori* SAP estimation for biased compensation. Moreover, our method uses another adaptive parameter to control the trade-off between speech distortion and residual noise for suppressing the estimated noise in highly nonstationary and various noisy environments. The performance of the new algorithm was evaluated by measuring the segment SNR and the ISDM. We showed that the proposed algorithm was generally superior to conventional methods, reducing both residual noise and speech distortion in nonstationary and noisy environments. In the future, we plan to evaluate its possible application in preprocessing for signal processing area.

## Competing Interests

The authors declare no competing interests.

## Acknowledgments

This research was supported by NRF (2013R1A1A2012536).

## References

- [1] S. Lee, C. Lim, and J.-H. Chang, "A new a priori SNR estimator based on multiple linear regression technique for speech enhancement," *Digital Signal Processing*, vol. 30, no. 7, pp. 154–164, 2014.
- [2] S. Lee and J. Chang, "On using multivariate polynomial regression model with spectral difference for statistical model-based speech enhancement," *Journal of Systems Architecture*, In press.
- [3] M. Berouti, M. Schwartz, and J. Makhoul, "Enhancement of speech corrupted by acoustic noise," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 208–211, Washington, DC, USA, April 1979.
- [4] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 32, no. 6, pp. 1109–1121, 1984.
- [5] Y. Hu, *Subspace and multitaper methods for speech enhancement [Ph.D. dissertation]*, University of Texas at Dallas, Richardson, Tex, USA, 2003.
- [6] O. Cappe, "Elimination of the musical noise phenomenon with the Ephraim and Malah noise suppressor," *IEEE Transactions on Speech and Audio Processing*, vol. 2, no. 2, pp. 345–349, 1994.
- [7] R. Martin, "Noise power spectral density estimation based on optimal smoothing and minimum statistics," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 5, pp. 504–512, 2001.
- [8] S. Rangachari and P. C. Loizou, "A noise-estimation algorithm for highly non-stationary environments," *Speech Communication*, vol. 48, no. 2, pp. 220–231, 2006.
- [9] I. Cohen, "Optimal speech enhancement under signal presence uncertainty using log-spectral amplitude estimator," *IEEE Signal Processing Letters*, vol. 9, no. 4, pp. 113–116, 2002.
- [10] I. Cohen, "Noise spectrum estimation in adverse environments: improved minima controlled recursive averaging," *IEEE Transactions on Speech and Audio Processing*, vol. 11, no. 5, pp. 466–475, 2003.
- [11] L. Lin, W. H. Holmes, and E. Ambikairajah, "Adaptive noise estimation algorithm for speech enhancement," *Electronics Letters*, vol. 39, no. 9, pp. 754–755, 2003.
- [12] S. J. Lee and S. H. Kim, "Noise estimation based on standard deviation and sigmoid function using a posteriori signal to noise ratio in nonstationary noisy environments," *International Journal of Control, Automation, and Systems*, vol. 6, no. 6, pp. 818–827, 2008.
- [13] R. Martin, "Bias compensation methods for minimum statistics noise power spectral density estimation," *Signal Processing*, vol. 86, no. 6, pp. 1215–1229, 2006.
- [14] D. Mauler and R. Martin, "Improved reproduction of stops in noise reduction systems with adaptive windows and nonstationarity detection," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, Article ID 469480, 17 pages, 2009.
- [15] P. C. Loizou, *Speech Enhancement: Theory and Practice*, CRC Press, Boca Raton, Fla, USA, 2007.
- [16] J. R. Deller, J. H. L. Hansen, and J. G. Proakis, *Discrete Time Processing of Speech Signals*, IEEE Press, New York, NY, USA, 1999.
- [17] Y.-S. Park and J.-H. Chang, "A probabilistic combination method of minimum statistics and soft decision for robust noise power estimation in speech enhancement," *IEEE Signal Processing Letters*, vol. 15, pp. 95–98, 2008.

## Research Article

# QoS and QoE Aware N-Screen Multicast Service

**Ghulam Sarwar, Farman Ullah, and Sungchang Lee**

*Department of Electronics and Information Engineering, Korea Aerospace University, Deogyang-gu, Goyang-si, Gyeonggi-do 412-791, Republic of Korea*

Correspondence should be addressed to Sungchang Lee; [sclee@kau.ac.kr](mailto:sclee@kau.ac.kr)

Received 17 December 2015; Accepted 29 March 2016

Academic Editor: Marco Anisetti

Copyright © 2016 Ghulam Sarwar et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The paper focuses on ensuring the quality-of-service (QoS) and quality-of-experience (QoE) requirements of users having heterogeneous devices in a multicast session. QoS parameters such as bit rate, delays, and packet losses are good indicators for optimizing network services but fall short in characterizing user perception (QoE). In N-Screen service, the users have different devices with heterogeneous attributes like screen size, resolution, and access network interface, and the users have different QoE on N-Screen devices with the same QoS parameters. We formulate the objective function of the N-Screen multicast grouping to ensure the minimum user's QoE with smaller bandwidth requirement. We propose a dynamic user reassignment scheme to maintain and satisfy the QoE by adapting the user's membership to the varying network conditions. The proposed schemes combine the available bandwidth and multimedia visual quality to ensure the QoS and QoE. In the network architecture, we introduce the functions of the QoS and QoE aware multicast group management and the estimation schemes for the QoS and QoE parameters. The simulation results show that the proposed multicast service ensures the network QoS and guarantees the QoE of users in the varying network conditions.

## 1. Introduction

Multimedia streaming is one of the fastest growing applications and the last decade alone has witnessed a rapid development in the complexity of smart devices and communication technologies. The interactive advertisement bureau report states that about 69% of users have 4-Screens [1] active and about 90% of the users' interaction is through Multi-Screen [2]. N-Screen, also termed as Multi-Screen, has different attributes such as network interfaces, media codecs, screen resolutions, computing, and storage. N-Screen service enables users to access and share the content from the same source through heterogeneous devices from any location and at any time [3, 4]. The ubiquity of communication technologies and smart devices demands that users should be able to access their subscribed services through any of the available networks such as 3G, 4G, LTE, and WiFi, as well as through any device such as smart phones, laptops, desktops, and tablets. N-Screen multimedia service demands extensive bandwidth to ensure the quality of service (QoS). A lack of bandwidth could cause higher end-to-end delay, video jitter, and packet losses. Cisco forecasts that by 2019,

80% of the global IP data will be video traffic. Moreover, the heterogeneity of N-Screen devices raises new questions about the multimedia QoS and user's experience on different terminals [4, 5]. The challenges that come with a rapid increase in mobile video traffic and ensuring multimedia QoS and user's quality of experience (QoE) on mobile devices can be addressed by either boosting the network capacity or by offloading traffic from the core and backhaul networks [6]. This paper presents a novel N-Screen multimedia multicast scheme to minimize the traffic and ensure that the QoE requirements of all group users on their N-Screen devices are met.

The huge growth rates of multimedia services utilize higher bandwidth and cause congestion in the core network. IP multicast [7] is a technique designed to conserve the network bandwidth by sending a single stream to multiple hosts simultaneously which helps in reducing the core network traffic. The multicast stream receiving nodes can have different capabilities and QoS and QoE requirements. IP multicast synchronizes the content transmission rate to the worst channel conditions among the multicast users that cause the traditional bandwidth fairness problem. To improve fairness

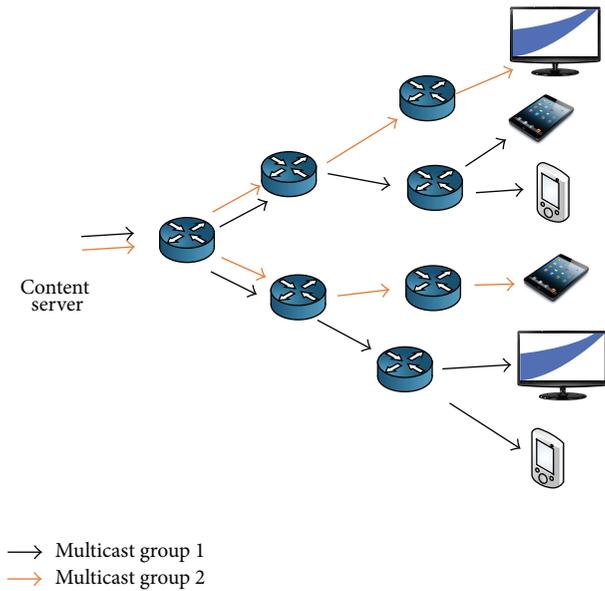


FIGURE 1: N-Screen aware multicast service.

among the multicast users, the destination set grouping (DSG) [8, 9] schemes proposed group users based on their available bandwidth (ABW). The existing DSG schemes do not consider the QoE requirements and different capabilities of N-Screen devices and cannot ensure the QoE requirements to all users as the QoE may be different on different devices in the same network conditions. Moreover, N-Screen devices may have the same QoE in different network conditions. The QoS requirements of N-Screen users can be ensured by defining different QoS traffic classes and serving high-quality classes with priority as in Differentiated Services [10] and IEEE802.11e [11]. Alternatively, the scalable video coding [12, 13] adapts the data transmission rate to the varying network conditions to handle the distribution of content to heterogeneous devices. These models require network transcoding to handle the different QoS classes and varying network conditions. This paper focuses on ensuring the satisfaction of QoS and QoE requirements of N-Screen multicast users without requiring in-network multimedia adaptation as that in SVC or layered video coding. We formulate the objective function of the N-Screen partitioning problem to ensure and maximize the users' QoE. The suggested scheme assigns users to the appropriate visual quality based multicast groups (MGs) at the source as shown in Figure 1.

From a network perspective, QoS is characterized by bandwidth constraints, delays, packet losses, and jitters that affect video quality. Tavakoli et al. [14] investigated the effect of the chunk/packet size and the quality switching amplitude on the QoS and QoE. Although QoS can give valuable insights about the network conditions, it alone is often inadequate for providing a reliable estimation of the video quality perceived by the end user, which is known as QoE [15]. QoS and QoE together characterize the quality of multimedia services that can be provided by the current network conditions and perceived by users of N-Screen devices. The QoS is a performance contract that can be measured either qualitatively

or quantitatively between the user and service provider. The QoE is a qualitative metric which can be measured with Motion-Based Video Integrity Evaluation Index (MOVIE) [16] or mean opinion score (MOS) [17]. In this study, we adopt the MOS metric of ITU-T G.1070 [8, 18] standard which considers the N-Screen attributes and network conditions information for QoE estimation. Andrew Catellier et al. [19] stated that users have different multimedia QoS and QoE on different devices and that the perceived multimedia quality is not the same on 10' inch tablets and 21' inch LCD screens. Therefore, considering only the bit rate does not satisfy the QoE requirements of each user in the MG as the perceived visual quality is different for each user's N-Screen device. In this paper, we propose a bit rate and multimedia quality aware N-Screen multicast service to ensure that the QoS and QoE requirements of users in an N-Screen environment are met.

The paper is organized as follows: Section 2 describes the background and related works and Section 3 presents an overview of the proposed N-Screen multicast service. Section 4 explains the architecture for provisioning QoS and QoE aware N-Screen multicast service and schemes for available bandwidth estimation, as well as multimedia quality estimation. We presented the objective function of N-Screen MG management, dynamic user reassignment, and periodic MG management schemes in Section 5. Section 6 provides a discussion and a comparison of the proposed architecture simulation results with that of previously existing ones. And lastly, Section 7 concludes the paper.

## 2. Background and Related Works

In this section, we introduce the background knowledge and related works concerning N-Screen service, multicast service, and multimedia QoS and QoE before going into the details of the proposed N-Screen multicast service.

N-Screen service enables users to use multiple heterogeneous devices at any time and at any location. The N-Screen technology aims to provide seamless service continuity and synchronize the content over heterogeneous displays [20]. In reality, N-Screen service is the enabling of one-source multiuse (OSMU) with that of an adaptive source multidevice (ASMD). A recent work [3, 4] defined and classified the various services that can be provided on multiple devices, and also the handover mechanisms over various access networks and devices. In [21], the authors provided the implementation details of mobile DTV broadcasting for N-Screen smart devices on top of iOS and Android OS platforms. Kwon et al. [22] identified the different factors such as service quality, ease of use, and price fairness that affect the behavioral intent of using existing N-Screen service. A scalable framework is proposed in [23] to provide web-based N-Screen user interface management for different screen layouts. A seamless screen switching scheme without intervention of user is proposed considering the user's context extracted from the embedded sensors in the devices [24]. Mostly, these research works have focused on the definitions and architectures of N-Screen services. Moving forward, this paper focuses on the grouping of heterogeneous N-Screen devices for multicast-based data dissemination to alleviate the network

congestion that currently exists in the mobile and fixed networks.

The multicast service provides a one-to-many or many-to-many communication to reduce the overall core network load. The multicast schemes can be application layer multicast [25, 26], IP-layer multicast [27], and link layer multicast [28]. In the multicast environment, a user may have different channel conditions that arise due to interference, fading, obstacles, and distance from the base station. The intergroup proportional fairness and multicast proportional fairness schemes are proposed in [29] to enhance the system throughput by assigning appropriate data rates to MGs according to the channel conditions. The IP Multimedia Subsystems (IMS) support in the Universal Mobile Telecommunications System (UMTS) became available in the 3rd Generation Partnership Project (3GPP) Release 5 to support service continuity in heterogeneous networks. The Release 5 of 3GPP originally only supported the unicast transmission that functions to increase the IP-core network load [30]. However, the 3GPP Release 6 and later versions [31] introduced the multimedia broadcast and multicast service (MBMS) in IMS. Converged IMS-MBMS integrated frameworks have been proposed to support multicast/broadcast services and mobility in heterogeneous networks [30–32]. Despite being superior to previous models, the MBMS service still does not consider the different capabilities and QoE requirements of N-Screen devices.

QoS is a measure of service availability and transmission quality of a network. QoS management schemes focus on application layer [33], network layer [28], and MAC layer [34] mechanisms to adapt the session to network conditions. These schemes require accurate information of the network QoS parameters such as ABW, packet loss ratio (PLR), jitters, and average packet delay. The estimation of the ABW is one of the critical factors for managing QoS and network congestion. ABW can be estimated either as hop-to-hop or end-to-end. The individual hop bandwidth estimation tools are proposed in previous studies [35, 36], and the end-to-end bandwidth estimations are *Pathload* [37] and *pathChirp* [38]. The end-to-end ABW in the network is a time-varying process and can be defined as the minimum of nonutilized capacities of all the links in between. The bandwidth estimation can be categorized as (i) *passive*, which is estimated from the packet loss, delay, and congestion without intruding into the network, and (ii) *active*, which sends probe packets to estimate the bandwidth. The active probing method is further classified as probe gap or probe rate models. The Trains of Packet Pairs [39] sends out packet pairs that are well separated in time and uses linear regression to estimate the bandwidth. The *pathChirp* uses exponential time space between the probe packets to estimate the ABW, which can accurately capture the status of network's congestion. *Pathload* uses the one-way delay of periodic packets to estimate the ABW and states that when the stream rate is larger than the ABW, the one-way delay of the periodic packets shows an increasing trend. A Bayesian-based prediction model was proposed to predict instantaneous end-to-end bandwidth change and the prediction method of one-way bandwidth estimation at the receiver [40]. Ahuja et al. [41] proposed the network such as

2G and 3G selection based on available link in the multiaccess networks environment.

In the multimedia streaming, one of the important factors is the insurance of the user's QoE requirements. The multimedia QoE estimation models can be classified as media layer models (MLMs) or packet layer models (PLMs). MLMs use the multimedia content to estimate the visual quality while the PLMs use the packet and network information to predict the visual quality [42]. In a previous research [43], the authors categorized the multimedia QoE estimation models as Full reference, Reduced reference, and No reference models. Full reference models [16] require both the original stream and the received distorted stream, Reduced reference models require some parameters description of the original stream and the received distorted stream, and No reference models use only the distorted stream. ITU-T J.144 [44] and ITU-R BT.1683 [45] are the Full reference models of the perceived visual quality estimation in the Digital TV. Kusuma et al. [46] developed the Reduced reference model for the estimation of perceptual visual quality. No reference parametric packet layer models have been proposed or studied in the past [47–49]. ITU-T G.1070 [8] is the recommendation based on the PLM of QoE estimation for videophone multimedia streaming services.

The contribution and significance of this paper compared to the existing approaches can be summarized as follows. (1) We focus on ensuring the QoS and QoE requirements of users having heterogeneous devices in a multicast session. (2) We propose the N-Screen multicast user partitioning as an optimization problem to maximize the users' QoE using smaller users' bandwidth and provide an algorithm to manage the number of MGs. (3) We propose a dynamic user reassignment scheme to maintain the QoE by adapting the user's membership according to the varying network conditions. (4) To provide N-Screen multicast service, we introduce the schemes of QoS estimation, QoE estimation, and visual quality aware MG management in the network architecture.

### 3. Overview of the Proposed N-Screen Multicast Service

In the proposed N-Screen multicast service, we assume that N-Screen users specify their visual quality requirements (MOS) and the service provider side assigns the users to appropriate MGs that can meet users' requirement as shown in Figure 2. In the figure, the  $x$ -axis represents the content transmission rate of MGs at the source and  $y$ -axis represents the expected perceived visual quality on user device. In an N-Screen multicast environment, devices having different network conditions can be assigned to the same MG and vice versa. This is because the MOS depends not only on the bit rate and PLR, but also on the codec, frame rate, and screen resolution. We assume that the visual quality requirement of a user is less than or equal to the maximum visual quality that can be provided in the current network conditions on the user's device. The criteria for assigning a user to a MG and the goals of the N-Screen multicast service are briefly discussed as follows.

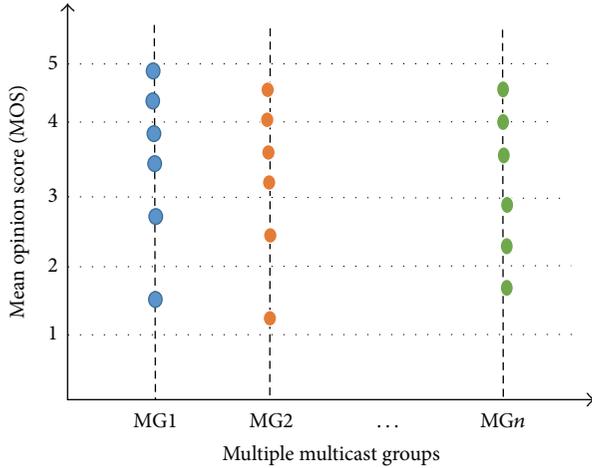


FIGURE 2: N-Screen, network, and visual quality aware user assignment to different multicast groups.

- (i) Assign user to one of the existing MGs that can provide the required visual quality on user's device.
- (ii) Assign user to the MG that provides maximum visual quality and requires smaller bandwidth as compared to other existing MGs.
- (iii) Create a new MG if none of the available MGs can meet user's requirement subject to the availability of system's resources.

#### 4. The Proposed QoS and QoE Aware N-Screen Multicast Service Architecture

Multicast disseminates the same content simultaneously to multiple receiving nodes to efficiently utilize the network resources. The main challenge of multimedia multicast service is providing the downlink data dissemination to N-Screen devices having different capabilities, QoS requirements, locations, and network conditions. Although QoS parameters such as available bandwidth, packet loss, and end-to-end delay of the underlying networks are good indicators for application service providers to optimize their services, these parameters do not reflect the perceived visual quality on the end user device. In N-Screen, users experience different visual quality on different N-Screens at the same QoS parameters due to the heterogeneous capabilities of the devices. We propose a novel N-Screen multicast service that focuses on ensuring the QoS and QoE requirements of N-Screen users at the service provider side by grouping users to appropriate MGs without requiring in-network multimedia adaptation as that in scalable video coding. In the following subsections, we introduce the proposed architecture, estimation of available bandwidth, and visual quality on N-Screen device.

*4.1. Proposed N-Screen Multicast Service Architecture.* To support N-Screen multicast service, we introduce the functions of N-Screen MG management, QoS and QoE estimation techniques, and N-Screen service repository in the service

provider network as shown in Figure 3. We briefly discuss the functions of each node of Figure 3 as follows.

*4.1.1. N-Screen Service Management System (NSMS).* To efficiently support N-Screen service across heterogeneous networks and devices, there should be a node that captures the entire network information. The NSMS is the central control unit that performs the registration and authorization of user's device and provides directory service. It maintains users' device list, IP addresses, and content lists of content servers.

*4.1.2. N-Screen Profile and Multicast Group Manager (PMGM).* After service registration with the NSMS, the user performs device registration with the PMGM. The PMGM maintains the current network conditions information and device attributes of each N-Screen device and assigns users to appropriate MGs. It gets the user's QoS parameters from the N-Screen content server (NCS) and uses them with N-Screen attributes to estimate the perceived visual quality. If the group membership of a device is changed due to changes in the network conditions, the PMGM informs the NCS to invite the user to the appropriate MG for better service quality.

*4.1.3. N-Screen Content Server (NCS).* NCS performs session management and QoS parameters estimation and sends the adapted contents to MGs. It also contains a bandwidth estimation module that sends probing packet trains to N-Screen devices to estimate the end-to-end ABW and PLR and reports them to the PMGM for further actions.

*4.2. QoS Parameters Estimation.* The multimedia services over the best-effort networks (i.e., Internet) are confronted with many challenges, and one of the most important factors is the ABW. The unpredictable and dynamic behavior of internet traffic adversely affects the multimedia quality. ABW and PLR are critical metrics to ensure the user's QoS requirements. We adapt and extend the concept proposed in previous works [50, 51] by using probe trains to estimate the ABW. Figure 4 shows the proposed ABW estimation scheme using the Poisson distribution for the interpacket gap and an adaptive number of probing packets in the trains to accurately capture the network's congestion.

We define the maximum and minimum number of packets in a train and adapt the number of packets between the maximum and minimum limits. The bandwidth estimation module of the NCS sends back-to-back probe packets and uses one-way delay (OWD) to estimate the ABW. We linearly increase the number of probe packets in the next train if the estimated  $OWD_{train}$  is greater than a predefined threshold. If the  $OWD_{train}$  is less than the threshold, we reduce the number of packets in the next trains. Let  $t_i$  be the time-stamp the probe packet  $i$  starts transmission with and let  $t'_i$  be its receiving time at the receiver. The OWD of the consecutive

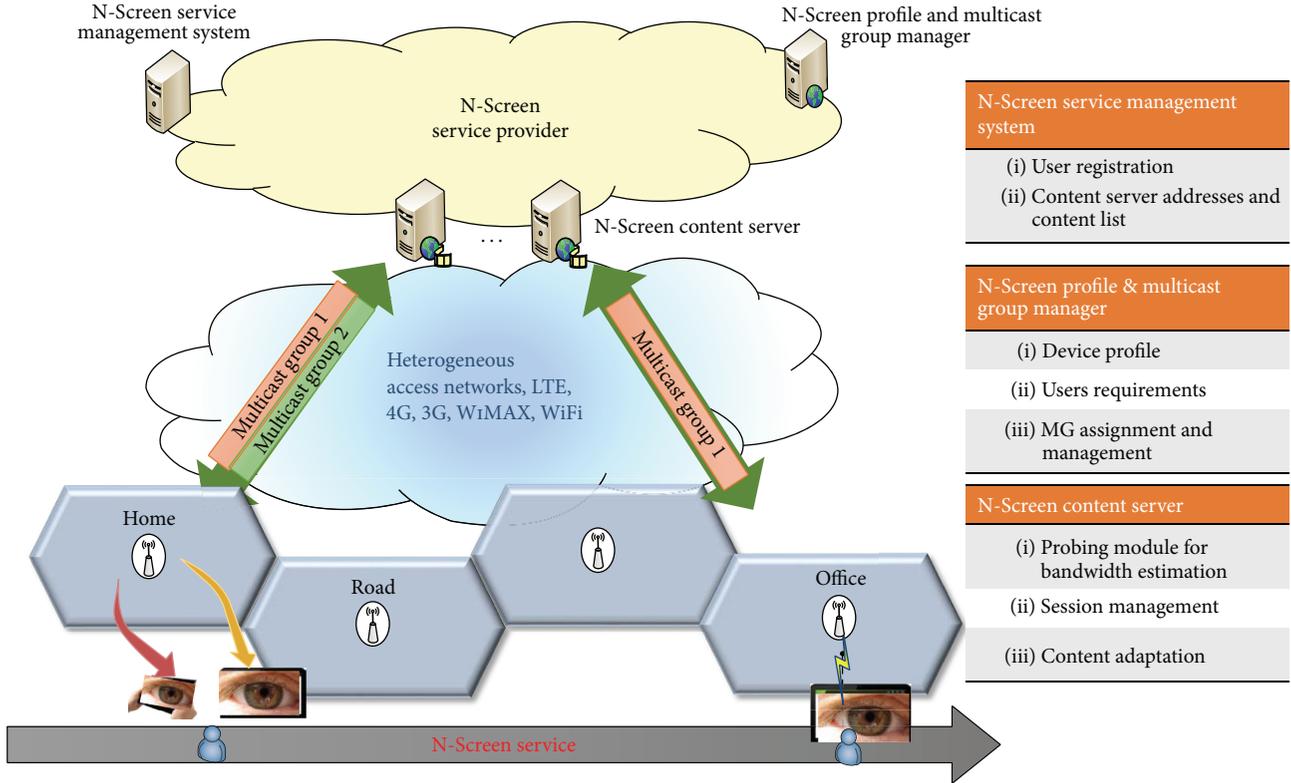


FIGURE 3: The proposed QoS and QoE aware N-Screen service architecture.

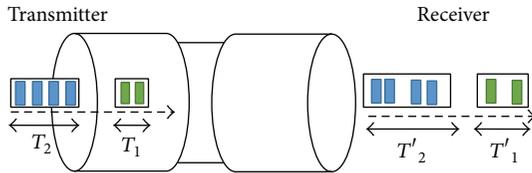


FIGURE 4: Available bandwidth estimation using probing train method.

probe packets in the probe train can be  $(t_{i+1} - t_i) - (t'_{i+1} - t'_i)$ . The  $OWD_{train}$  can be found by the following:

$$OWD_{train} = \frac{\sum_{i=1}^{N-1} (t_{i+1} - t_i) - (t'_{i+1} - t'_i)}{N - 1}, \quad (1)$$

where  $N$  is the number of probe packets in the train. The available bandwidth  $ABW$  for probe packet size  $S$  can be found by the following:

$$ABW = \frac{\sum_{i=1}^{N-1} S_i}{OWD_{train}}. \quad (2)$$

Packet loss is another important QoS parameter and occurs due to many factors such as queuing, user location, and channel noise in the wireless networks. The packet losses reduce the system throughput and link reliability and increase the transmission delay due to retransmissions. The multimedia quality is highly susceptible to the packet loss as real-time

multimedia services are mostly based on the unreliable User Datagram Protocol (UDP). The PLR is given by the following:

$$PLR = \frac{ExPackets - RePackets}{ExPackets}, \quad (3)$$

where  $ExPackets$  is the expected packets (transmitted packets) and  $RePackets$  is the actually received packets. Each N-Screen device reports the reception time-stamps of the packets and the  $ExPackets$  of the train to the bandwidth estimation modules to estimate the  $AWB$  and the  $PLR$  using (2) and (3).

**4.3. QoE (Visual Quality) Estimation.** QoS parameters give full insight about the network status but do not give any information about the multimedia quality perceived by the end user on his N-Screen device. QoE parameters estimate the perceived visual quality on user's device. The perceived visual quality depends on the capabilities of N-Screen devices and access network conditions. A user experiences different visual quality at different N-Screen devices even in the same network conditions. We adapt the concept of ITU-T G.1070 [8] to estimate the multimedia quality (MOS) on user's N-Screen device:

$$MOS = 1 + \left( v_3 - \frac{v_3}{1 + [ABW/v_4]^{v_5}} \right), \quad (4)$$

where  $MOS$  is the mean opinion score and its range is from 1 to 5. The variables  $v_3$ ,  $v_4$ , and  $v_5$  depend on the media code, frame rate, and screen size.

## 5. QoS and QoE Aware N-Screen Multicast Group Management

In this section, we introduce the QoS and QoE aware user's assignment and multicast group management. When a user requests the desired content, considering its QoS and QoE requirements and regarding the ABW, the PLR, and the expected MOS of the content on his current device, the MG management function decides whether to add the user to one of the existing MGs or a new MG. In a mobile network, the supported data rates and packet losses at the receiving nodes in a MG may be different due to different channel conditions and locations. In the paper, we approximate the PLR and ABW of a MG at the base node, that is, the node having highest losses and smallest ABW in the MG. We introduce the N-Screen group management, dynamic user reassignment, and periodic MG management schemes in the following subsections.

*5.1. N-Screen Group Management.* The objective of the proposed N-Screen grouping is to ensure and maximize the users' QoE using smaller bandwidth. The perceived visual quality in a MG is an increasing function of the content transmission rate (CTR) as long as the CTR is less than the ABW and is decreasing when the CTR exceeds the ABW due to the occurrence of PLR. The objective of the partitioning problem is to appropriately assign the users to MGs that ensure and maximize users' experience using smaller users' bandwidth. Let  $V(U, K)$  denote the session utility "V" of partitioning users  $1 \cdots U$  into  $K$  groups that maximize the user's experience using smaller bandwidth. We may express  $V$  as follows:

$$V(N, K) = \sum_{u=1}^N \sum_{k=1}^K \lambda_{u,k} * \frac{((r_u - r_k) * 5) / \text{MaxCap}}{\text{MOS}_u - \text{MOS}_{u,r_k}}, \quad (5)$$

such that

$$0 \leq \sum_{g=1}^G \text{BW}_g \leq \alpha T \leq T, \quad (6)$$

$$\text{MOS}_{u,k} \leq \text{RQ}_u, \quad (7)$$

$$\text{BW}_k \leq \min \{ \text{ABW}_i \}_{i=1}^{U_k}, \quad (8)$$

$$\forall K_i, K_j, i \neq j : K_i \cap K_j = 0, \quad (9)$$

$$1 \leq U_k \leq U_T, \quad (10)$$

$$1 \leq K \leq U_T, \quad (11)$$

$$\sum_{i=1}^K U_k = U_T, \quad (12)$$

where  $\lambda_{u,k} \in \{0, 1\}$  is the assignment of user  $u$  to MG  $k$ ,  $r_u$  is the available bandwidth of user  $u$ , is the CTR of MG  $k$ , and  $\text{MaxCap}$  is the maximum link's capacity. We scale the bit rate in (5) from 0 to 5 to make it comparable to the MOS parameter in the equation. The constraint in (6) states

that bandwidth assigned to MGs should not overwhelm the system threshold bandwidths which is 60% in the standard eMBMS specification. Equation (7) states that the visual quality provided by the MG "k" should satisfy the user's requirements. The constraint of (8) shows that the CTR of a MG "k" should be less than minimum ABW among all the users of the group. Notations section shows the symbols and notation descriptions used in these equations and this paper. The system performs user's assignment to the groups using Dynamic Programming Algorithm as in [52]. The group membership of a user is decided by using the following decision equation:

$$\lambda_{u,k} = \begin{cases} 1; & \text{if } \frac{((r_u - r_k) * 5) / \text{MaxCap}}{\text{MOS}_u - \text{MOS}_{u,r_k}} \geq \frac{((r_u - r_j) * 5) / \text{MaxCap}}{\text{MOS}_u - \text{MOS}_{u,r_j}} \\ & \text{for } j = 1, 2, \dots, K, j \neq k, \text{ subject to MOS constraints} \\ 0; & \text{otherwise.} \end{cases} \quad (13)$$

*5.2. Dynamic User Reassignment.* The perceived visual quality by N-Screen users from multicast groups varies due to the varying network conditions, and fixing a user's assignment to multicast groups cannot ensure the visual quality requirements of the user during the lifetime of the session. We suggest dynamic user reassignment to maintain the perceived visual quality and ensure the user's requirements. The dynamic user reassignment algorithm is shown in Figure 5. Considering the current network conditions, user's requirements, and N-Screen attributes, the proposed scheme monitors the visual quality deterioration. If the visual quality deterioration condition is satisfied, the suggested scheme searches for another existing MG that can provide better visual quality and may ensure user's requirement. The system invites the user to the appropriate MG to maintain the QoE requirements.

## 6. Evaluation Results

The section provides the comprehensive simulation results and comparison with other schemes. The results show that the proposed multicast service efficiently utilizes the network resources and provides the QoS and QoE endurance. We use the OMNET++ based simulation environment [53]. Figure 6 shows the simulation model to estimate the available bandwidth and multimedia visual quality. In the figure, S are the sources that generate the background traffic, R are the routers, AP are the wireless network base stations, and the N-Screen devices are four heterogeneous devices with each AP. Table 1 shows the simulation parameters used in the simulation and evaluation of the scheme.

Figure 7 shows the comparison of the bandwidth estimation schemes using the probe gap model, the linear increment in the number of probes, and the proposed Poisson distribution based interpacket gap model with an adaptive number of packets in the trains. The proposed scheme predicts fast bandwidth fluctuation as compared to the other schemes. The result clearly demonstrates that traditional methods

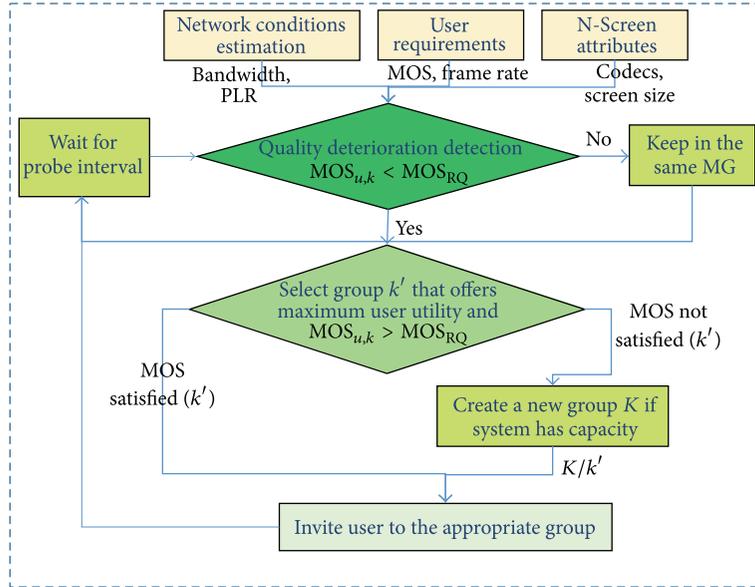


FIGURE 5: Algorithm of dynamic user reassignment.

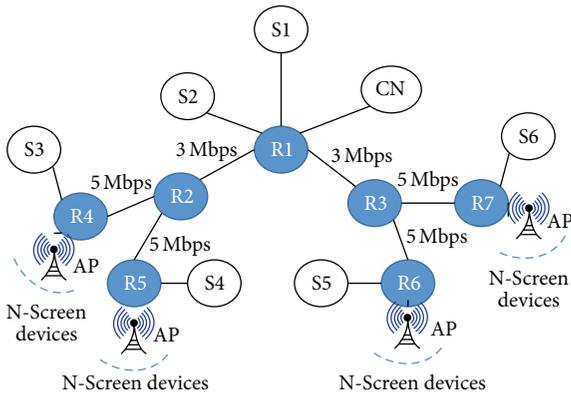


FIGURE 6: Simulation environment.

TABLE 1: Simulation setup parameters.

Parameter	Value
Probe packet size	1000 bytes
Bottleneck links capacity	3 Mbps
Other links	5 Mbps
Frame rate	30 fps
Minimum probe packets in train	2
Maximum probe packets in train	5

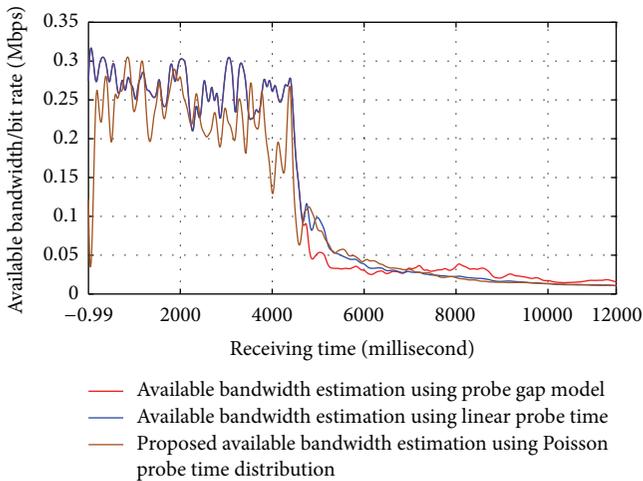


FIGURE 7: Comparison of the available bandwidth estimation schemes.

do not have the caliber to abruptly detect the changes in the available bandwidth. Figure 7 shows the scenario of normal background traffic and at 44960 ms, we increase the background traffic to detect the change in the bandwidth. The proposed scheme detects the change about 97 ms earlier than the other schemes. Since the proposed scheme can detect the changes in the available bandwidth earlier than schemes so the source can adjust its bit rate in advance. This helps to decrease the congestion, reduce the packet losses, and improve the user’s QoE.

The perceived visual quality strongly depends on the network conditions and N-Screen attributes as shown in Figure 8(a). The codecs with high compression ratio such as H.264 and higher transmission rate can provide better visual quality in good channel conditions, but as the PLR increase less, a codec with less compression ratio such as MPEG-4 and less transmission rate offers better visual quality than H.264. The system assigns a user to the MG that offers better visual quality on an N-Screen device in the current network conditions. Due to the unpredictable nature of background internet traffic, we suggest dynamic MG reassignment to maintain and ensure the required QoE as depicted in Figure 8(b).

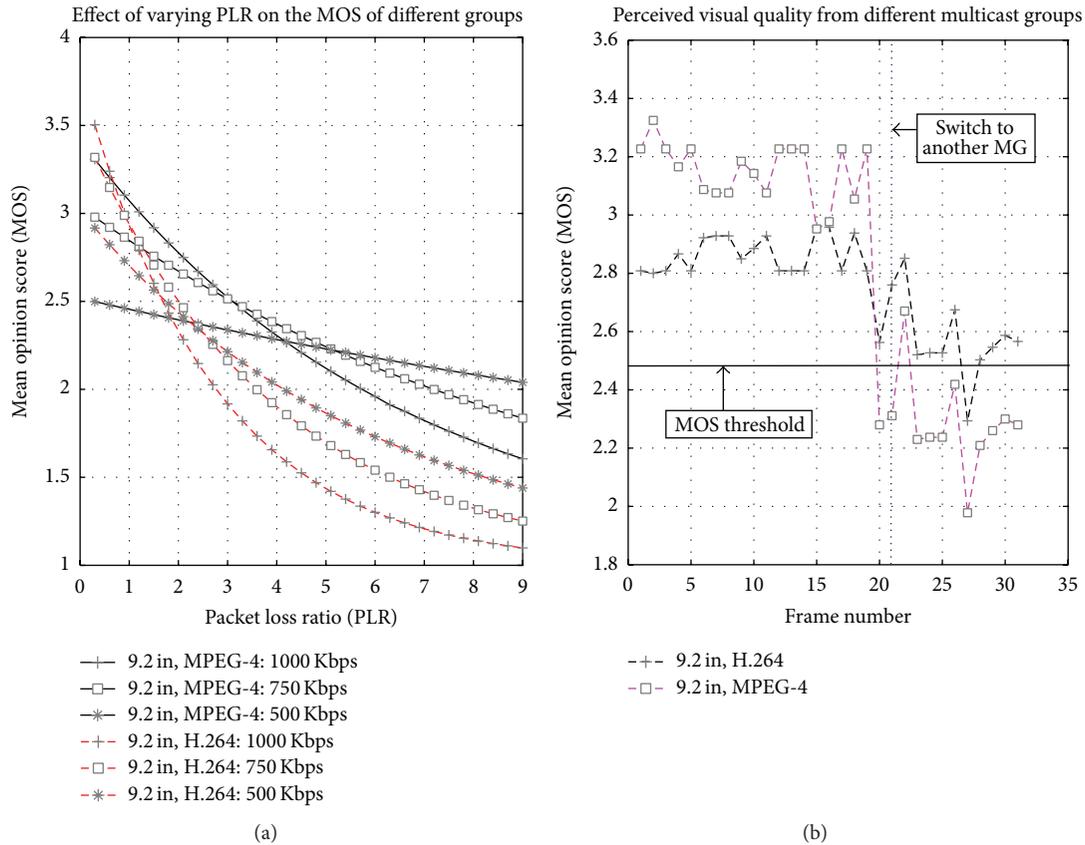


FIGURE 8: Effect of dynamic group assignment on visual quality.

TABLE 2: Comparison of N-Screen multicast with unicast and multicast schemes.

Device ID	Screen size	QoE requirements	ABW (Kbps)	PLR	Unicast MOS	Bit rate based multicast scheme		Proposed N-Screen multicast	
						MOS	Grouping	MOS	Grouping
D-1	2.1"	MOS > 2.5	300	0.6	3.48	3		3	
D-2	4.2"	MOS > 2	500	0	3.48	2.11	Transmit at base rate (200) to all users but cannot ensure the visual quality requirements of all users	2.11	MG1: base rate (200)
D-3	2.1"	MOS > 2.5	200	0	3.18	3.18		3.18	
D-4	9.2"	MOS > 2	500	0	2.25	1.73		2.18	MG2: base rate (450)
D-5	4.2"	MOS > 2	450	2	2.46	1.77		2.46	

Table 2 shows a comparison of the proposed QoS and QoE aware N-Screen multicast service with the traditional bit rate based multicast and unicast schemes. The bit rate based multicast scheme groups devices in a single multicast session although the QoE requirements of D-4 and D-5 are not fulfilled. Unicast based transmission offers high visual quality to N-Screen users by utilizing the user's ABW at the cost of network resources. The proposed scheme divides the users into MGs to ensure that visual quality requirements of all N-Screen users are met. Figure 9 compares the efficiency of the proposed N-Screen multicast service in terms of perceived visual quality with that of traditional unicast and multicast schemes. In the experiment, we vary the number of users in the session.

## 7. Conclusion

In the paper, we proposed QoS and QoE aware adaptive mapping of N-Screen devices to multicast groups (MGs) at the application layer to ensure the visual quality requirements of N-Screen devices in varying network conditions without requiring network transcoding. In N-Screen service, users have heterogeneous devices that have different attributes, and the users have different QoE on different devices with the same QoS. We proposed the architecture, management functions, and the parameters estimation schemes for providing N-Screen multicast service. The proposed multicast service ensures the minimum QoS and QoE requirements of users in the varying network conditions. Furthermore, the proposed

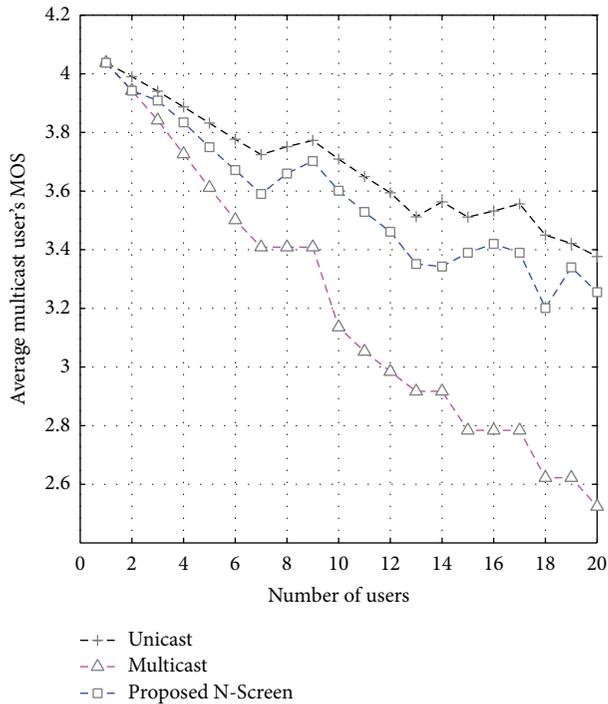


FIGURE 9: Comparison of the proposed N-Screen multicast with unicast and multicast schemes.

available bandwidth estimation scheme can abruptly detect the changes in the network conditions. The proposed N-Screen multicast scheme can efficiently utilize the network resources and offers comparable visual quality as that by a unicast session.

## Notations

OWD:	One-way delay for a packet
$BW_{av}$ :	End-to-end available bandwidth
S:	Probe packet size
PLR:	Packet loss ratio
MOS:	Mean opinion score (multimedia quality range: 1–5)
G:	Number of multicast groups
U:	Number of unicast users in the system
$ U_g $ :	Number of multicast users in the multicast group
$\alpha$ :	Total bandwidth of the system
$BW_g$ :	Fraction of bandwidth assigned to multicast services
$BW_i$ :	Bandwidth assigned to multicast group $i$
$U_g - U_k = 0$ :	Users will belong to one multicast group for the same service.

## Competing Interests

The authors declare that there is no conflict of interests regarding the publications of this paper.

## References

- [1] "The Multi-Screen Marketer: Interactive Advertising Bureau Report," 2012, [http://www.iab.net/media/file/The\\_Multiscreen-Marketer.pdf](http://www.iab.net/media/file/The_Multiscreen-Marketer.pdf).
- [2] Australian Multi-Screen Report, 2014, <http://www.nielsen.com/content/dam/niensenglobal/au/docs/reports/australian-multi-screen-report-q32014.pdf>.
- [3] C. Yoon, T. Urn, and H. Lee, "Classification of N-screen services and its standardization," in *Proceedings of the 14th International Conference on Advanced Communication Technology (ICACT '12)*, February 2012.
- [4] F. Ullah, G. Sarwar, H. Lee, W. Ryu, and S. Lee, "Control framework and services scenarios of provisioning n-screen services in interactive digital signage," *Tehnicki Vjesnik*, vol. 21, no. 6, pp. 1321–1328, 2014.
- [5] J. Kim, T.-W. Um, W. Ryu, B. S. Lee, and M. Hahn, "IPTV systems, standards and architectures: part II—heterogeneous networks and terminal-aware QoS/QoE-guaranteed mobile IPTV service," *IEEE Communications Magazine*, vol. 46, no. 5, pp. 110–117, 2008.
- [6] S. Spagna, M. Liebsch, R. Baldessari et al., "Design principles of an operator-owned highly distributed content delivery network," *IEEE Communications Magazine*, vol. 51, no. 4, pp. 132–140, 2013.
- [7] S. Deering, "Host extensions for IP multicasting," RFC 1112, 1989, <http://www.ietf.org/rfc/rfc1112.txt>.
- [8] ITU-T G, "Opinion model for video-telephony applications," Recommendations 1070, 2007.
- [9] S. Y. Cheung, M. H. Ammar, and X. Li, "On the use of destination set grouping to improve fairness in multicast video distribution," in *Proceedings of the 15th Annual Joint Conference of the IEEE Computer Societies, Networking the Next Generation (INFOCOM '96)*, vol. 2, pp. 553–560, San Francisco, Calif, USA, March 1996.
- [10] K. Nichols, S. Blake, F. Baker, and D. Black, "Definition of the differentiated services field (DS Field) in the IPv4 and IPv6 headers," RFC RFC2474, 1998.
- [11] D. Kaur, K. Kaur, and V. Arora, "QoS in WLAN using IEEE802.11e (Survey of QoS in MAC layer protocols)," in *Proceedings of the 2nd International Conference on Advanced Computing and Communication Technologies (ACCT '12)*, pp. 468–473, Rohtak, India, January 2012.
- [12] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 9, pp. 1103–1120, 2007.
- [13] L. Wang, Z. Yang, L. Xu, and Y. Yang, "NCVCS: network-coding-based video conference system for mobile devices in multicast networks," *Ad Hoc Networks*, 2016.
- [14] S. Tavakoli, K. Brunnström, J. Gutiérrez, and N. García, "Quality of experience of adaptive video streaming: investigation in service parameters and subjective quality assessment methodology," *Signal Processing: Image Communication*, vol. 39, pp. 432–443, 2015.
- [15] F. Tommasi, V. De Luca, and C. Melle, "Packet losses and objective video quality metrics in H.264 video streaming," *Journal of Visual Communication and Image Representation*, vol. 27, pp. 7–27, 2015.
- [16] K. Seshadrinathan and A. C. Bovik, "Motion tuned spatio-temporal quality assessment of natural videos," *IEEE Transactions on Image Processing*, vol. 19, no. 2, pp. 335–350, 2010.

- [17] K. Seshadrinathan, R. Soundararajan, A. C. Bovik, and L. K. Cormack, "Study of subjective and objective quality assessment of video," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1427–1441, 2010.
- [18] Excel, "Video Quality Analysis Software G. 1070 Revision 2007 + G.107 Revision 2006," ITU-T G 1070, Excel, 2007, <http://www.e-model.org/>.
- [19] A. Catellier, M. Pinson, W. Ingram, and A. Webster, "Impact of mobile devices and usage location on perceived multimedia quality," in *Proceedings of the 4th International Workshop on Quality of Multimedia Experience (QoMEX '12)*, pp. 39–44, Victoria, Australia, July 2012.
- [20] K. Ha, K. Kang, and J. Lee, "N-screen service using I/O virtualization technology," in *Proceedings of the International Conference on Information and Communication Technology Convergence (ICTC '10)*, pp. 525–526, IEEE, Jeju, South Korea, November 2010.
- [21] S.-H. Park, Y.-S. Park, S.-B. Yu, and J.-R. Choi, "Implementation of ATSC mobile DTV broadcasting for N-screen smart devices," in *Proceedings of the IEEE International Conference on Consumer Electronics (ICCE '12)*, pp. 331–332, Las Vegas, Nev, USA, January 2012.
- [22] B.-R. Kwon, S.-H. Ryu, and Y.-G. Kim, "Identifying factors affecting behavioral intent of potential and existing N-screen service users," *ETRI Journal*, vol. 37, no. 2, pp. 417–427, 2015.
- [23] Y. Bae, B. Oh, and J. Park, "Adaptive transformation for a scalable user interface framework supporting multi-screen services," in *Ubiquitous Information Technologies and Applications*, vol. 280 of *Lecture Notes in Electrical Engineering*, pp. 425–432, Springer, Berlin, Germany, 2014.
- [24] Y. T. Hong, Y. S. Hwang, S. H. Nam et al., "A seamless screen switching scheme for N-screen services based on context awareness," *Wireless Personal Communications*, 2016.
- [25] S. Banarjee, B. Bhattacharjee, and C. Kommareddy, "Scalable application layer multicast," in *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM '02)*, pp. 205–217, Pittsburgh, Pa, USA, August 2002.
- [26] J. H. Wang, J. Cai, J. Lu, K. Yin, and J. Yang, "Solving multicast problem in cloud networks using overlay routing," *Computer Communications*, vol. 70, pp. 1–14, 2015.
- [27] S. E. Deering, "Host extensions for IP multicasting," 1988.
- [28] E. Cerqueira, L. Veloso, A. Neto, M. Curado, E. Monteiro, and P. Mendes, "Mobility management for multi-user sessions in next generation wireless systems," *Computer Communications*, vol. 31, no. 5, pp. 915–934, 2008.
- [29] H. Won, H. Cai, K. Guo et al., "Multicast scheduling in cellular data networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 9, pp. 4540–4549, 2009.
- [30] A. Ikram, M. Zafar, N. Baker, and R. Chiang, "IMS-MBMS convergence for next generation mobile networks," in *Proceedings of the IEEE International Conference on Next Generation Mobile Applications, Services and Technologies (NGMAST '07)*, pp. 49–56, Cardiff, Wales, September 2007.
- [31] 3GPP TS 23.246 v7.2.0, Multimedia Broadcast/Multicast Service (MBMS), 2004.
- [32] M. Knappmeyer, B. Ricks, R. Tönjes, and A. Al-Hezmi, "Advanced multicast and broadcast content distribution in mobile cellular networks," in *Proceedings of the 50th Annual IEEE Global Telecommunications Conference (GLOBECOM '07)*, pp. 2097–2101, Washington, DC, USA, November 2007.
- [33] Y. C. Yee, K. N. Choong, A. L. Y. Low, and S. W. Tan, "SIP-based proactive and adaptive mobility management framework for heterogeneous networks," *Journal of Network and Computer Applications*, vol. 31, no. 4, pp. 771–792, 2008.
- [34] M. Oche, R. M. Noor, and J. I. Aghinya, "Network centric QoS performance evaluation of IPTV transmission quality over VANETs," *Computer Communications*, vol. 61, pp. 34–47, 2015.
- [35] V. Jacobson, *Pathchar: A Tool to Infer Characteristics of Internet Paths*, 1997.
- [36] B. A. Mah, pchar: a tool for measuring internet path characteristics, 2001, <http://www.kitchenlab.org/www/bmah/Software/pchar/>.
- [37] M. Jain and C. Dovrolis, "Pathload: a measurement tool for end-to-end available bandwidth," in *Proceedings of the Passive and Active Measurements (PAM '02)*, pp. 14–25, March 2002.
- [38] V. Ribeiro, R. Riedi, R. G. Baraniuk, J. Navratil, and L. Cottrell, "Pathchirp: efficient available bandwidth estimation for network paths," in *Proceedings of the Passive and Active Measurement Workshop (PAM '03)*, April 2003.
- [39] P. Selin, K. Hasegawa, and H. Obara, "Available bandwidth measurement technique using impulsive packet probing for monitoring end-to-end service quality on the Internet," in *Proceedings of the 17th Asia Pacific Conference on Communications (APCC '11)*, pp. 518–523, Sabah, Malaysia, October 2011.
- [40] A. Javadtalab, M. Semsarzadeh, A. Khanchi, S. Shirmohammadi, and A. Yassine, "Continuous one-way detection of available bandwidth changes for video streaming over best-effort networks," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 1, pp. 190–203, 2015.
- [41] K. Ahuja, B. Singh, and R. Khanna, "Network selection based on available link bandwidth in multi-access networks," *Communications and Networks*, vol. 2, no. 1, pp. 15–23, 2016.
- [42] A. Takahashi, D. Hands, and V. Barriac, "Standardization activities in the ITU for a QoE assessment of IPTV," *IEEE Communications Magazine*, vol. 46, no. 2, pp. 78–84, 2008.
- [43] M. Carnec, P. Le Callet, and D. Barba, "Full reference and reduced reference metrics for image quality assessment," in *Proceedings of the 7th International Symposium on Signal Processing and Its Applications (ISSPA '03)*, pp. 477–480, July 2003.
- [44] ITU-T, "Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference," Recommendation J.144, 2004.
- [45] ITU, "Objective perceptual video quality measurement techniques for standard definition digital broadcast television in the presence of a full reference," ITU-R BT.1683, 2004.
- [46] T. M. Kusuma, H. Zepernick, and M. Caldera, "On the development of a reduced-reference perceptual image quality metric," in *Proceedings of the Systems Communications*, pp. 178–184, Montreal, Canada, August 2005.
- [47] A. R. Reibman, V. A. Vaishampayan, and Y. Sermadevi, "Quality monitoring of video over a packet network," *IEEE Transactions on Multimedia*, vol. 6, no. 2, pp. 327–334, 2004.
- [48] J. Joskowicz and J. C. L. Ardao, "A parametric model for perceptual video quality estimation," *Telecommunication Systems*, vol. 49, no. 1, pp. 49–62, 2012.
- [49] M. Cheon, S.-J. Kim, C.-B. Chae, and J.-S. Lee, "Quality assessment of mobile videos," in *Visual Signal Quality Assessment*, C. Deng, L. Ma, W. Lin, and K. N. Ngan, Eds., pp. 99–127, Springer, 2015.
- [50] F. Ullah, G. Sarwar, and S. Lee, "A network and visual quality aware N-screen content recommender system using joint

- matrix factorization,” *The Scientific World Journal*, vol. 2014, Article ID 806517, 13 pages, 2014.
- [51] D. Kim and K. Chung, “A network-aware quality adaptation scheme for device collaboration service in home networks,” *IEEE Transactions on Consumer Electronics*, vol. 58, no. 2, pp. 374–381, 2012.
- [52] Y. R. Yang, M. S. Kim, and S. S. Lam, “Optimal partitioning of multicast receivers,” in *Proceedings of the International Conference on Network Protocols*, pp. 129–140, Osaka, Japan, November 2000.
- [53] OMNET++ based Simulator for Wireless and Mobile Networks, <http://mixim.sourceforge.net/>.

## Research Article

# Energy Optimization for Outdoor Activity Recognition

**Mehdi Boukhechba,<sup>1</sup> Abdenour Bouzouane,<sup>1</sup> Bruno Bouchard,<sup>1</sup>  
Charles Guin-Vallerand,<sup>2</sup> and Sylvain Giroux<sup>3</sup>**

<sup>1</sup>LIARA Laboratory, University of Quebec at Chicoutimi, Chicoutimi, QC, Canada G7H 2B1

<sup>2</sup>LICEF Research Center, Tele-Universite of Quebec, Quebec, QC, Canada G1K 9H6

<sup>3</sup>DOMUS Laboratory, University of Sherbrooke, Sherbrooke, QC, Canada J1K 2R1

Correspondence should be addressed to Mehdi Boukhechba; [m.boukhechba@esi.dz](mailto:m.boukhechba@esi.dz)

Received 20 December 2015; Accepted 30 March 2016

Academic Editor: Chehri Abdellah

Copyright © 2016 Mehdi Boukhechba et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The mobile phone is no longer only a communication device, but also a powerful environmental sensing unit that can monitor a user's ambient context. Mobile users take their devices with them everywhere which increases the availability of persons' traces. Extracting and analyzing knowledge from these traces represent a strong support for several applications domains, ranging from traffic management to advertisement and social studies. However, the limited battery capacity of mobile devices represents a big hurdle for context detection, no matter how useful the service may be. We present a novel approach to online recognizing users' outdoor activities without depleting the mobile resources. We associate the places visited by individuals during their movements with meaningful human activities using a novel algorithm that clusters incrementally user's moves into different types of activities. To optimize the battery consumption, the algorithm behaves variably on the basis of users' behaviors and the remaining battery level. Studies using real GPS records from two big datasets demonstrate that the proposal is effective and is capable of inferring human activities without draining the phone resources.

## 1. Introduction

Human's activity recognition has represented an active topic of research for several decades. However, only in the recent years, with the increasing availability and facilities of collecting movement datasets from GSM or GPS equipped devices or even network wireless technologies like WI-FI [1] and RFID [2], have we had the possibility to study users' activities from their movement traces.

Mobile tracking devices, for example, phones and navigation systems, sense the movement of persons represented by positioning records that capture geolocation, time, and a number of other attributes. Sensing is based on a collection of information related to the achieved activity from raw sensor data (GPS, Wi-Fi, RFID, Bluetooth signals, microphone, camera, accelerometers, magnetometers, gyroscopes, barometers, proximity sensors, etc.) to extract a pertinent information about the current activity. In context awareness systems, the challenge relies on developing applications that sense and

react to environmental changes to provide a value-added user experience.

As such, the mobile phone is no longer only a communication device, but also a powerful environmental sensing unit that can monitor a user's ambient context, both unobtrusively and in real time. Ambient sensing [3–5] has become a primary input for a new class of mobile services like activity recognition. In fact, real-time recognition of users' activities offers the possibility to understand what they are doing at the present moment and estimates their actions in the future. This context awareness property makes this field a major piece that provides services to a range of application domains such as real-time traffic monitoring [6, 7], social networking [8], and cognitive assistance.

Smartphones are ubiquitous and becoming more and more sophisticated, with ever-growing computing, networking, and sensing powers. This has been changing the landscape of individuals' daily life and has opened the doors for many interesting data mining applications like activity

recognition. Sensors are the source for raw data collection in activity recognition [9]; however, the limited battery capacity of mobile devices represents a big hurdle for the quality and the continuity of the service. The embedded sensors in the mobile devices are major sources of power consumption. Hence, excessive power consumption may become a major obstacle to broader acceptance of context-aware mobile applications, no matter how useful the service may be.

Context information can be related to the user environment, but also to the device itself. Since smartphones are battery-powered, in an ideal scenario, the application will self-adapt and adjust its behavior according to the current battery status of the device. It is in this context that our research lies; we will try to propose a battery-aware activity recognition solution in order to preserve mobiles' life battery.

We bring a novelty to the activity recognition field via three points:

- (1) We propose a novel self-adaptive clustering approach that adjusts the computational complexity of the algorithm on the basis of the remaining battery level. The goal is to prevent the massive draining of the mobile resources in order to capture users' movements for the longest time possible. Our mining method is based on a new version of online  $K$ -means, where we propose a temporal data window that is characterized by a variable size on the basis of a person's moving behavior and his phones' remaining resources.
- (2) The majority of related works are based on the classification of historical records of people's trajectories using density based approach (see, e.g., [10]), where they try to identify the most visited place with post-treatment processes, for example, end of the day and every week. These methods fail in their ability to deal with the places less visited by people but important in their trajectories, for example, cemetery and airport, and cannot be used in fields like assistance where we need a real-time access to a person's activity. That is why we propose a new online solution which answers those difficulties.
- (3) This work handles not only stationary behaviors but also moving activities like shopping. We introduce the speed and the variance of the orientation of people's trajectories as a new variable in our system for this purpose.

In this paper, we will demonstrate an innovative method to online switch raw GPS data to meaningful human activities. This method uses only a mobile device without network or historical record requirements, while consuming a minimum of mobile resources. The presented approach aims to enrich people's movements, represented in real time during their travel trajectories, with semantic information about the visited places. Our method is based on the real-time recognition of points of interest, "POI" (a place of interest is an (urban) georeferenced object where a person may carry out a specific activity), in users' trajectories. This service increases the use of this contribution, contrasting from economic uses like traffic management, public transportation, commercials,

and advertising, to more serious uses like security, police, and risk management.

The following sections detail our contribution: Section 2 briefly reviews related work; Section 3 presents our approach in terms of three major components, that is, trajectory classification, spatial recognition, and activity discovery; Section 4 describes the experimentation by highlighting two dimensions: accuracy and power saving. Finally, conclusion and future focus, as well as the expected contributions, are summarized in Section 5.

## 2. Related Works

The emerging concept of activity recognition using mobile devices is a new topic for trajectory data analysis; however, research community's efforts are increasing day by day to carry clear definitions and common understandings [6]. We will highlight two parts of related works, works on activity recognition and semantic trajectories and battery-aware works.

*2.1. Activity Recognition and Semantic Trajectory Works.* Kang et al. in [1] utilized the access point MAC address of a WI-FI network to capture location data on a campus. They developed a time-based clustering algorithm to "extract places" taking advantage of the continuity of the WI-FI positioning. A new place is found when the distance of the new locations from the previous place is beyond a threshold and when the new locations span a significant time threshold. This algorithm is simple and works in an incremental way on mobile devices. However, the algorithm does not consider the reoccurrence of readings at the same location. More simply, each time it discovers a place, it is a "different" place. This also makes it difficult to discover places that are visited with high frequency but short-dwell time. Moreover, this method requires continuous location data collection with very fine intervals and, thus, large storage. Another shortage of this work is linked to the labeling of the discovered places; since it is made manually by authors, the work needs to be improved by providing an automatic way of labeling activities.

A clustering method called CB-SMOT (Clustering-Based Stops and Moves of Trajectories) is proposed in [10]; it is a clustering method based on the speed variation of the trajectory; it is used to infer semantic information from trajectories. Firstly, this method evaluates the trajectory sample points and generates clusters in places where the trajectory speed is lower than a given threshold for a minimal amount of time. In a second step, the method matches the clusters with a set of relevant geographic places defined by the user. However, the shortage in this solution is that data processing is nonincremental; it is not possible to analyze in real time users mobility using this solution. Moreover, as the capture of users' activities is based on stops, this solution fails to recognize moving activities, such as walking in a park.

CityVoyager, presented in [11], is a recommendation system designed for mobile devices, which recommends shops to users based on data analyzed from their past location history. Their system applies location data to the item-based

collaborative filtering algorithm, an algorithm used in many online recommendation systems, by transforming location data history into a list that contains the names of each user's frequently visited shops and rating values indicating the user's preference for each shop. Authors track the visited shops by the loss of GPS signal. Nevertheless, it is known that GPS signals are frequently lost in urban areas even when the user is outdoors; these situations increase the possibilities of false detections. Furthermore, it is also known that it is possible to visit a shop without losing GPS signal, in a case where the shop does not include many obstacles hindering the diffusion of the GPS signal; this situation makes these shops unrecognizable by this work.

Authors claim to propose an approach designed for mobile phones. However, there is no adaptation noted to support this demanding environment. For instance, finding frequented shops requires heavy manipulation of the historical records of the users' visited shops. Authors seem to neglect the limited mobile's resources, since there is no support for limited battery life and there is no effort perceived to online detect and find the frequent shops.

In [12], an algorithm is proposed to associate each stop in a user's trajectory to a list of possible visited places and each of these places is associated with a probability. Finally, depending on the kinds of activities associated with the identified place, the trajectory is classified into probable trajectory behavior. In this work, authors assume the moving object is a person that travels using transportation means associated with a traceable (GPS) device (car, bus, metro, or train). The person gets off the transportation means to walk to the final destination. During this time interval, the person is not traceable; accordingly, authors have used probabilities to find the visited place. This work uses numerous thresholds that are set manually as minimum duration of an activity. Nevertheless, since these parameters may depend on user profiles, this work may be ineffective on large datasets that contain several profiles.

While developing a rich body of work for managing moving objects, the research community has shown a little interest in the online recognition of POI in users' trajectories. The majority of related works are based on the classification of historical records excluding problems linked to mobile's performance, like battery life and low-computational capacities.

Moreover, nearly all approaches are based on the detection of stops within people's movements, neglecting activities with movements, and only a minority of these studies try to automatically identify the background geographic information, since generally we request a set of relevant geographic places defined manually by the user.

The majority of outdoor activity recognition approaches use a fixed activity's minimum duration threshold that represents the minimum time a user has to spend in the POI (place of interest) to be declared as visited place. This threshold prevents false activity detection, such as traffic jams. However, previously fixing this threshold will increase error probability because when set to a small value, it will increase the number of false activities, like passing by a POI; setting it to a high value will miss detection of some short-dwell activities, such as buying cigarettes at the

convenience store. Thus, we propose a novel approach to online recognize users' visited places; our algorithm will be totally unsupervised, which operates without any beforehand fixed threshold.

*2.2. Battery-Aware Works.* The problem of power management on mobile devices has been well-explored. Viredaz et al. [13] surveyed many fundamental but effective methods for saving power on handheld devices. These methods concern a range of phone components such as processor, memory, display screen, audio system, and wireless networking. It has been suggested from the architectural point of view that the system hardware should be designed as a collection of interconnected building blocks that could function independently to enable independent power management.

In [14], authors proposed a dynamic frequency/voltage scaling (DVS) to reduce power consumption by configuring the processor based on the requirements of the executing applications. It is recognized as the basis of numerous energy management solutions [15]. DVS exploits the fact that dynamic power consumption is a strictly convex function of the CPU speed and attempts to save energy by reducing the supply voltage and frequency at runtime. Hence, the power-saving scheme should be fully customized for real-time power consumption situations and specific application requirements. However, these methods are more suitable for lower-level system design rather than application development.

Shih et al. [16] introduced a technique to increase battery lifetime of a personal digital assistant- (PDA-) based phone by reducing its idle power, the power a device consumes in a "standby" state. To do so, they essentially shut down the device and its wireless network card when the device is not being used, avoiding energy waste while not activated. The device is powered only when an incoming call is received or when the user needs to use the PDA for other purposes.

Authors in [17] have gone further by proposing a whole framework of Energy Efficient Mobile Sensing System (EEMSS) for automatic user-state recognition. The core component of EEMSS is a sensor management scheme for mobile devices that operates sensors hierarchically, by selectively turning on the minimum set of sensors to monitor a user state, and triggers new set of sensors if necessary to achieve state transition detection. Energy consumption is reduced by shutting down unnecessary sensors at any particular time.

As seen above, significant efforts have been undertaken for a wise use of mobile resources; however, the proposed techniques try to limit the calculation capacities to gain in battery life. Our perspective is different from this one. We propose a self-adaptive approach that changes dynamically the calculation capacities on the basis of battery life and user state. Our work will also stand out by the flexibility that it offers to users, since they have the freedom to choose the degree of austerity in the use of the battery.

### 3. Overview of the Approach

We assume the person is GPS traceable via a smartphone. Usually, a person's activities are divided into two

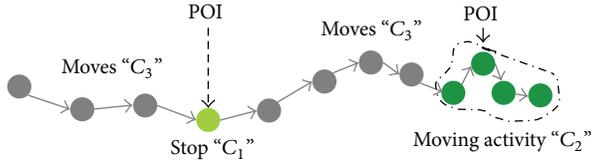


FIGURE 1: Relation between moves, stops, and activities with moving.

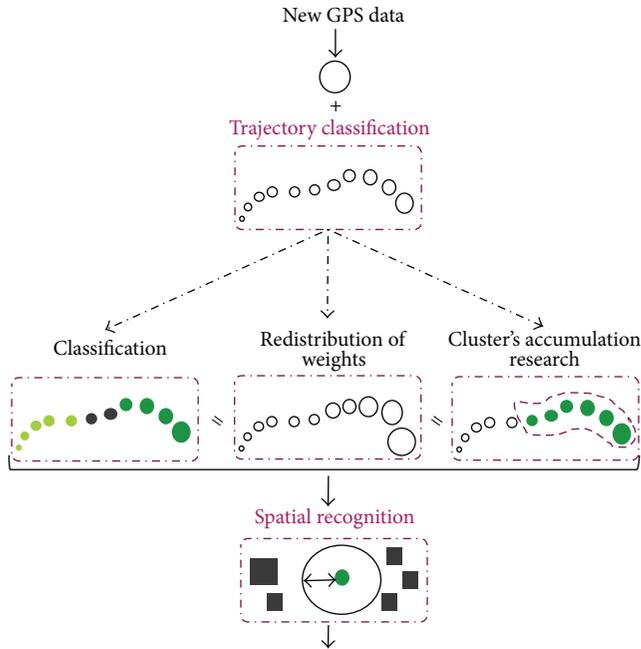


FIGURE 2: The overall process of our activity recognition approach.

behaviors: stationary and nonstationary, where the second one is also divided into two categories moving to reach a goal and moving to do a goal.

For example, working in the office is a stationary activity while going from the workplace to a shopping center is nonstationary activity; shopping itself is also a nonstationary activity, but the goal is to do shopping, so it is an activity with movements (see Figure 1). Based on these concepts we introduce 3 types of clusters:

- (1) Stop concept, represented by “ $C_1$ ”, characterizes stationary activities.
- (2) Activity with movements “ $C_2$ ” is nonstationary activities that require movement over a time interval.
- (3) Moves, represented by “ $C_3$ ,” are a set of actions that aim to move from a POI to another.

To deal with all these concepts, we present in Figure 2 the overall approach of our activity recognition mechanism.

In the first step, we introduced a real-time classification method based on  $K$ -means to classify every new position data according to the three families (stops, moves, and activity with movements). At the same time, we observe the accumulation of types of clusters, such that, after a certain threshold of the same cluster’s accumulation, we conclude

that the person is probably doing something interesting. For the second step, we summarize the accumulated clusters to a probable POI and we begin a geospatial research for the closest and the most meaningful geographical entity. If the research process succeeds, we determine this point as a POI. The third step is to assign the POI to an activity, like a museum to a tourism activity and gym to a sports activity.

**3.1. Step 1: Trajectory Classification.** The aim of this step is to incrementally classify the continuous users’ positions into different kinds of activities. The most recent sequence of GPS collection is stored in a temporal window called TW.

**Definition 1.** GPS collection is an assembly of GPS points  $P = \{p_1, p_2, \dots, p_n\}$ . Each GPS point  $p_i \in P$  contains a latitude ( $p_i.lat$ ), a longitude ( $p_i.Lngt$ ), a timestamp ( $p_i.T$ ), a speed ( $p_i.S$ ), and a bearing ( $p_i.B$ ). We add to this information the variance of bearing ( $p_i.V$ ) of the last  $L_{TW}$  GPS points where  $L_{TW}$  is the size of our temporal window TW and the weight ( $p_i.W$ ) represents the importance of the point  $p_i$  according to the time generation.

**Definition 2.** Temporal window TW is a subgroup of a GPS collection with a variable length  $L_{TW}$ . In fact TW contains all  $p_i$  with not null weight  $p_i.W$  (see Figure 3 that represents the relation between a GPS collection and TW; every point in the GPS collection is a record row from the database).

Once new GPS data is received, we achieve three parallel processes like below.

**3.1.1. Process 1: Classification.** At the arrival of new GPS data,  $p_n$  is stored in TW. Classification process is not launched on every data arrival but after a threshold called  $T_{min}$  that will be exposed in process 3. We classify  $p_i$  in TW using online  $K$ -means; we consider two variables, speed  $p_i.S$  and variance of bearing  $p_i.V$ .

The variance of bearing  $p_i.V$  is calculated using this formula:  $p_i.V = \sum(p_i - \bar{p}_i)/l$ , where  $\bar{p}_i = \sum p_i/L_{TW}$ ; this calculation is made before recording the new  $p_i$  in TW and it represents the variance of user’s orientation in the last  $L_{TW}$   $p_i$ .

Stops behaviors are characterized by a very low  $p_i.V$  and  $p_i.S$ ; however, moves behaviors are characterized by high  $p_i.S$  and low  $p_i.V$  because a person’s movements using transportation tend to be in a quick and straight manner. Moving activities are branded by a low  $p_i.S$  and high  $p_i.V$  (see Figure 4) since these activities are pedestrian actions which generally require a frequent shift of orientation like visiting a museum, shopping in a mall, or walking in a zoo. As said previously, these three types of clusters represent three families of activities and not three activities, each family containing a set of activities depending on the user’s visited places.

Note that it is known that when a user stops somewhere, his positions during this stop may vary in the surroundings of the stop due to the positioning error of the tracking system. Consequently, in our algorithm, we automatically put

Location ID	Latitude	Longitude	Time	Speed	Bearing	Variance	Weight
601	48.4210453	-71.0572413	Mon Nov10..	0.74	76.5	49.7	0.61
602	48.4210453	-71.0572532	Mon Nov10..	0.79	76.5	88.36	0.73
603	48.4210453	-71.0572567	Mon Nov10..	0.24	76.5	115.7	1.0

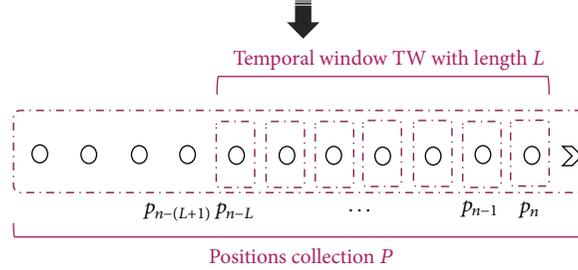
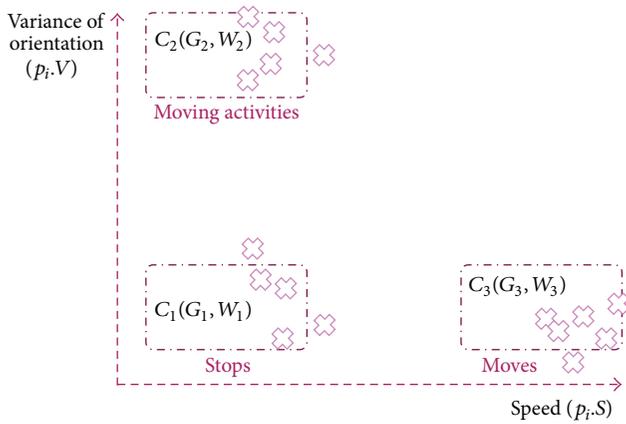
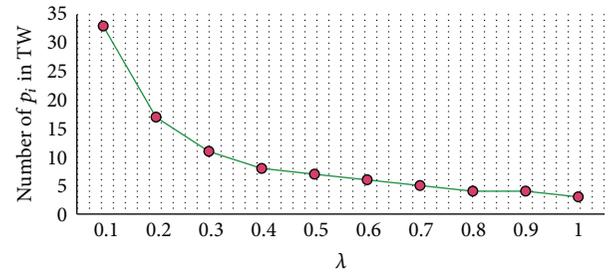
FIGURE 3: The relation between the GPS collection  $P$  and TW.

FIGURE 4: Inferring activity type using speed and the variance of orientation.

the bearing  $p_i.B \rightarrow 0$  when we detect that the speed  $p_i.S \rightarrow 0$  to avoid any error linked to this situation.

**3.1.2. Process 2: Distribution of Weights  $p_i.W$ .** Every  $p_i$  has a weight which determines the degree of resemblance of  $p_i$ 's class to the current activity. Our methodology is inspired from the work in [18]; the weight of each data point decreases exponentially with the time  $t$  via a fading function  $w_i = f(t) = 2^{-\lambda t}$ , where  $\lambda > 0$ . The exponentially fading function is widely used in temporal applications where it is desirable to gradually discount the history of the past behaviors. The parameter  $\lambda$  is called exponential decay constant; the higher the value of  $\lambda$ , the lower the importance of the historical data compared to more recent data. The overall weight of the data stream is a constant  $W = \sum_{t=0}^{t_c} 2^{-\lambda t} = 1/(1 - 2^{-\lambda})$  where  $t_c$  ( $t_c \rightarrow \infty$ ) is the current time.  $\lambda$  can also be seen as the determining parameter of TW's length. When  $\lambda$  approaches 1, TW shrinks to its smallest size. Inversely, when  $\lambda$  approaches 0, TW spreads to its maximum size (see Figure 5).

Unlike [18], we chose a variable value of  $\lambda$  between 0 and 1 depending on two parameters, the remaining battery level " $\beta$ " and the disorder of data " $E$ " that already exists in TW (class

FIGURE 5: The impact of varying  $\lambda$  on TW's length.

of every  $p_i$ ); the disorder is introduced in order to have an idea of the user instant state; for instance, if we have enough information on the user instant activity, there is no need to use a lot of data since the decision about the instant activity is already made; however, when the user state is difficult to be recognized, we will need more data to be able to detect the user activity; this ability to make a decision is called "disorder of data." Thus, these parameters are defined in the following definitions.

**Definition 3.** The remaining battery level is represented by  $\beta = \beta_r/\beta_{th}$ , where  $\beta_r$  is the real remaining battery level.  $\beta_{th}$  represents a battery threshold specified by the user from which the algorithm starts to minimize calculation (see Figure 6). This user-defined parameter adds some flexibility to the application. For instance, if the user prefers to keep good precision even if it drains the mobile battery,  $\beta_{th}$  should take a small value (e.g., 10%). Conversely, when the user aims for a good preservation of the mobile battery,  $\beta_{th}$  should be given a higher value of around 90% or 100%. Again, when the user knows that he will not spend much time outside, he can adjust  $\beta_{th}$  to a small value to promote the precision and vice versa.

**Definition 4.** The disorder of data  $E$  represents the quality of TW's data that has a link with the ability to make a decision. When one is sure that an activity is performed, one needs less

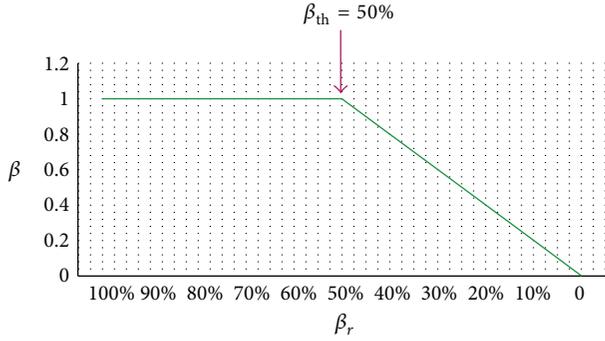


FIGURE 6: An example of  $\beta$  variation using a threshold  $\beta_{th} = 50\%$ .

data to make a decision so  $E \rightarrow 1$ , and  $E \rightarrow 0$  when one has problems finding out what type of activity is executed. In other terms, when the user behavior is unpredictable (because the data in TW is heterogeneous) one says that there is a disorder in his data ( $E \rightarrow 1$ ). Contrariwise, when one sees that the user behavior is stable, one says that there is a stability in the user data ( $E \rightarrow 0$ ).

The disorder of data is often calculated using the entropy, which is a measure of unpredictability of information content, or, in other terms, it measures the homogeneity in a set of information. Consequently,  $E$  is calculated using the entropy of Shannon [19];  $E$  of the new  $p_n$  is calculated using the entropy of the old data set in TW. The value of  $E$  is calculated as follows:

$$E = 1 - H_2(p_i) = 1 - \sum_{i=0}^{n-1} p_i \log_2 p_i, \quad (1)$$

where  $p_i = w_i/W$ .

After introducing the remaining battery level  $\beta$  and the data disorder  $E$ , we propose a relation between the two parameters to calculate  $\lambda$ ; we put

$$\lambda = 1 - \beta + E\beta. \quad (2)$$

The equation demonstration is shown below.

*Demonstration.* Our reflection starts from the following logical rules:

- (1) We shrink the TW's length to reduce the number of clustered points (in order to reduce battery consumption) when the battery is low or when we have some certitude about the user's activity, in other terms, stability in the user's behavior characterized by a low disorder of data.
- (2) We spread the length of TW when we need numerous position points to make a decision (high disorder in the user's TW data) as long as we have enough battery level.

TABLE 1: Mathematical representation of the user, battery, and TW states.

Characteristic	Interpretation
Battery high	$\beta \rightarrow 100\%$
Battery low	$\beta \rightarrow 0\%$
High disorder	$E \rightarrow 0$
Low disorder	$E \rightarrow 1$
TW spreads	$\lambda \rightarrow 0$
TW shrinks	$\lambda \rightarrow 1$

TABLE 2: Parsing  $\beta$ ,  $E$ , and  $\lambda$  to Boolean parameters.

Interpretation	Boolean parameters
$\beta \rightarrow 100\%$	$\beta$ true
$\beta \rightarrow 0\%$	$\beta$ false
$E \rightarrow 1$	$E$ true
$E \rightarrow 0$	$E$ false
$\lambda \rightarrow 1$	$\lambda$ true
$\lambda \rightarrow 0$	$\lambda$ false

TABLE 3: The truth table of  $\beta$ ,  $E$ , and  $\lambda$ .

$\beta$	$E$	$\lambda$
True	False	False
True	True	True
False	True	True
False	False	True

These conditions are sited as follows:

$$\begin{aligned}
 &\text{Battery high} + \text{high disorder} \rightarrow \text{TW spreads}, \\
 &\text{Battery high} + \text{low disorder} \rightarrow \text{TW shrinks}, \\
 &\text{Battery low} + \text{high disorder} \rightarrow \text{TW shrinks}, \\
 &\text{Battery low} + \text{low disorder} \rightarrow \text{TW shrinks}.
 \end{aligned} \quad (3)$$

If we parse the characteristics mentioned in (3) into a mathematical representation using the relations presented in Table 1, then, (3) will be written as follows:

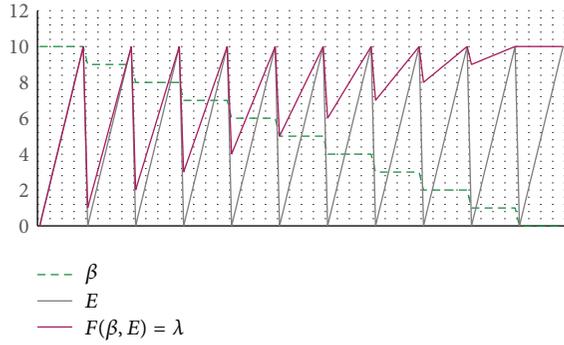
$$\begin{aligned}
 &\beta \rightarrow 100\% + E \rightarrow 0 \implies \lambda \rightarrow 0, \\
 &\beta \rightarrow 100\% + E \rightarrow 1 \implies \lambda \rightarrow 1, \\
 &\beta \rightarrow 0\% + E \rightarrow 0 \implies \lambda \rightarrow 1, \\
 &\beta \rightarrow 0\% + E \rightarrow 1 \implies \lambda \rightarrow 1.
 \end{aligned} \quad (4)$$

Let  $\beta$ ,  $E$ , and  $\lambda$  be Boolean parameters that take values as mentioned in Table 2.

The logical rules in (4) can be illustrated in the truth table, Table 3.

We notice that the truth table, Table 3, is that one of the Boolean implication function, where

$$\lambda \equiv F(\beta, E) \equiv \beta \wedge \neg E \equiv \beta \xrightarrow{\text{boolean}} E. \quad (5)$$

FIGURE 7:  $\lambda$  variation on the basis of  $\beta$  and  $E$ .

In Boolean logic, the truth values of variables may only be 0 or 1; however, our parameters  $\beta$ ,  $E$ , and  $\lambda$  take continuous values between 0 and 1. For instance,  $\beta = 1$  means full battery and  $\beta = 0$  means empty battery. Yet, in real life, there are many states between full and empty states. As such, the fuzzy logic was introduced [20] which is a form of many-valued logic in which the truth values of variables may be any real number between 0 and 1. Fuzzy logic has been extended to handle the concept of partial truth, where the truth value may range between completely true and completely false. In this context, our equation is called fuzzy implication:

$$\lambda \equiv F(\beta, E) \equiv \beta \xrightarrow{\text{fuzzy}} E. \quad (6)$$

Using the inequality of Reichenbach [21], the fuzzy implication can be written as

$$\lambda = F(\beta, E) = 1 - \beta + \beta E. \quad (7)$$

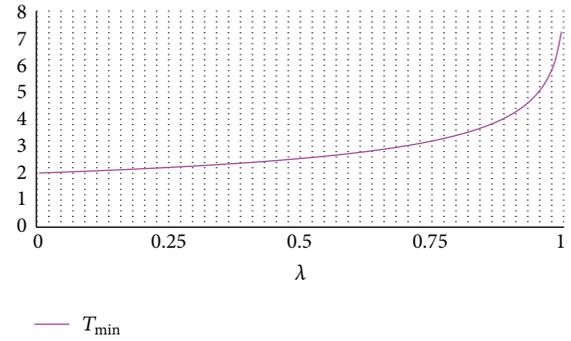
This equation has been tested in Figure 7 where we varied the two parameters  $\beta$  and  $E$  to observe the sensitivity of  $\lambda$  (the length of TW) to these changes.

In Figure 7, we notice that the sensitivity of TW's size (the value of  $\lambda$ ) is linked to the battery level  $\beta$  and disorder  $E$ ; however, when  $\beta$  begins to drop, the size of TW continuously ignores the disorder  $E$  until it reaches a total ignorance when  $\beta = 0$  (see Figure 7 when  $\beta = 0$ ).

After having distributed the weights of each  $p_i$  in TW and calculated the size of TW on the basis of the battery level  $\beta$  and disorder  $E$ , it is time to see if the user is performing some interesting activity.

**3.1.3. Process 3: Cluster's Accumulation Search.** The algorithm recognizes that someone is doing an activity if the weight of its cluster  $W_j$  exceeds a value  $\mu$ , where  $\mu = W/K$  with  $K$  representing the number of clusters (activities) used by  $K$ -means to classify TW; in our case  $K = 3$  because we are trying to identify three families of clusters: stops; moves; and moving activities.

The most important question is “when do we search for a cluster accumulation?” To minimize the use of device resources, it is recommended to handle this step carefully. The research process is not launched on every data arrival  $T$  but after a time called  $T_{\min}$  in which it is expected to have an activity.

FIGURE 8: The relation between  $\lambda$  and  $T_{\min}$ .

**Proposition 5.**  $T_{\min}$  is the time from which  $w_i = f(t) = 2^{-\lambda t}$  reaches  $\mu$ ; this is verified in this following condition  $2^{-\lambda T_{\min}} \mu + 1 = \mu$ ; after development  $T_{\min} = (1/\lambda) \log(\mu/(\mu - 1))$ , where  $\lambda = 1 - \beta + \beta E$ .

Consequently, on every  $T_{\min}$  we check if there is any activity that its weight  $W_j$  exceeds  $\mu$ ; if found, we summarize the points  $p_i$  to one point  $C_j(G_j, W_j)$ , where  $W_j = \sum_{i=0}^n w_{ij}$  and  $G_j = \sum p_{ij}/n$ ,  $n$  is the number of points in this cluster, and  $p_{ij}$  represents the points  $p_i$  in the cluster  $j$ . After that we move to the spatial recognition of the summarized point  $C_j$ .

As said previously,  $\lambda$  is not static. It varies between 0 and 1 depending on the disorder of data in TW and the remaining battery level. From the relation between  $\lambda$  and  $T_{\min}$  in Figure 8, we note that  $T_{\min}$  is also affected by these parameters. When  $\lambda$  is near 1, TW will contain a minimum number of position points (see Figure 5) for the reason that the battery is low or there is some stability in the user's behavior (e.g., staying at home); in this case, there is no need to process the calculations on every step; otherwise, this useless calculation depletes the battery. Consequently,  $T_{\min}$  will take a maximum value (see Figure 8).

Conversely, when  $\lambda$  is near 0, TW will contain a maximum number of position points (see Figure 5) for the reason that the battery is well charged and there is a big disorder in TW which troubles the users' activity identification. In this case the algorithm will process the calculations on every position-point arrival to quickly determine which type of activity the user is performing. Accordingly,  $T_{\min}$  will take a small value (see Figure 8).

Algorithm 1 achieves two processes; the first is to store every GPS data  $p_i$  when it arrives; the second is performed every  $T_{\min}$  to reduce the calculation. First step is to classify every point  $p_i$  in TW and then update the value of  $\lambda$  depending on the disorder of the activities types in TW and the remaining battery level. After calculating the center of gravity  $C_j$  of each cluster and the threshold  $\mu$ , we start to search for an accumulation of a cluster that is verified by the condition  $\max(W_j) > \mu$ ; we use the max of clusters' weight to avoid the case where two clusters exceed  $\mu$  in the same time; if condition is verified we start to recognize the geographic environment of  $C_j$ . Finally, we update the value of  $T_{\min}$  that will determine the next repetition of the process.

```

Input:
A GPS point  $p_i$ ;
Output:
The activity of the person;
(1) For each  $T$ 
(2)   Store  $p_i$  in TW;
(3) End for each
(4) For each  $T_{\min}$ 
(5)   Classify every point in TW;
(6)   Update  $\lambda$ ;
(7)   Update the centers of clusters  $C_j(G_j, W_j)$ ;
(8)   Calculate the threshold  $\mu$ ;
(9)   //Cluster's accumulation search
(10)  If ( $\max(W_j) > \mu$ ) then
(11)    POI = spatial recognition ( $G_j$ );
(12)    Activity = activity discovery (POI);
(13)  End if
(14) Update  $T_{\min}$ ;
(15) Return Activity
(16) End for each

```

ALGORITHM 1: Trajectory classification.

After searching for a cluster's accumulation, we move to Step 2, spatial recognition of  $C_j$ .

**3.2. Step 2: Spatial Recognition.** This task aims further to understand the movement behavior of users, in terms of more semantically meaningful POI. In this step, we are going to transform a cluster of GPS points to an expressive human activity. First we are going to search for the nearest geographic entities to this cluster; thereafter, we are going to associate an activity type to the entity found.

We search for the closest and the most significant geographical feature compared to  $C_j$ ; it is performed using a spatial query in a spatial database. Our database is powered by OSM (OpenStreetMap (OSM) is a community-driven project aimed at producing a high-quality detailed map of the world; OSM datasets are freely available under the Open Database license terms) and stored in the users' device for local use. This technique aims to discard networks use by offering offline services that will save a user's money and a phone's battery.

In fact, many mobile GIS solutions [22, 23] offer an offline version to avoid the constraints linked to the use of networks. For instance, Esri company [24] (one of the world's biggest GIS companies) has provided a whole runtime called "ArcGIS Runtime mobile SDK" [25] to support offline mapping; it includes map viewing, interaction, editing, and routing while fully disconnected from wireless.

Consequently, offline processing of geographic data has become easier. However, the limit of this technique is the size of the stored data, since geographic data may cause a massive use of the storage capacity; for instance, the size of the OSM planet database is over 666 GB [26]. The usual solution to face this problem is to download a specific database to each user depending on the cities where they live (e.g., the size of the whole New York region is only 140 MB [27] including all

geodata types such as roads, highways, buildings, and public parks); this can be done manually by the user, since he can define his area of interest while being online by choosing the area to download and panning/zooming to the area of capture. The second way is to download the geographic database automatically by the application by detecting the user's daily staying area; for instance, if the application detects that the user is living in Montreal city, it will wait until being online to download the OSM data of Montreal and its surroundings.

Many techniques to extract geographic data from OSM exist; the easiest way is to download and extract it from the OSM website. There are various web services that provide data extracts for a geographic area. For example, GeoFabrik is a company which specializes in working with OpenStreetMap. They provide a variety of free extracts in Shapefile (the Shapefile format is a geospatial vector data format for geographic information system (GIS) software) and raw OSM format on their download website (<http://download.geofabrik.de/>). After downloading raw OSM data and storing it in a spatial database, we use Algorithm 2 to search for the nearest spatial feature to  $C_j$ .

The querying process to search for the closest geographic entity is a famous technique based on identifying a proximity buffer; proximity analysis determines the relationship between selected geographic elements by identifying the locations of other elements within a specified distance (50 m in our case). Creating buffer zone regions is the most common method used in proximity analysis; for more details about the querying method please refer to the paper described in [28].

If the spatial recognition step finds a geographic entity in the neighborhood of the positions' cluster we can declare that the user has visited this place.

**3.3. Step 3: Activity Discovery.** Activity discovery is based on the exploitation of tags in OSM. OpenStreetMap data adhere to a simple XML schema with three basic elements: (i) nodes, that is, single geospatial points; (ii) ways, intended as ordered sequences of nodes; (iii) relations, grouping multiple nodes and/or ways. Each element includes a unique identification code, latitude and longitude coordinates, versioning information, and optional general-purpose informative tags. A tag is a key-value pair of Unicode strings of up to 255 characters. Each tag describes a specific feature of a spatial data element (see [29] for more details). This step permits having a human meaningful semantic information about the user visited place.

Tags are written in OSM documentation as key = value. The key describes a broad class of features (e.g., highways or names). The value details the specific feature that was generally classified by the key; for example, the geographic entity that contains a tag "building = apartment" represents a building arranged into individual dwellings, often on separate floors.

We assume that each tag represents an activity and each activity belongs to an activity family (see Figure 9).

Activities are organized in a taxonomy which generalizes the kinds of activities of interest for the movement analysis.

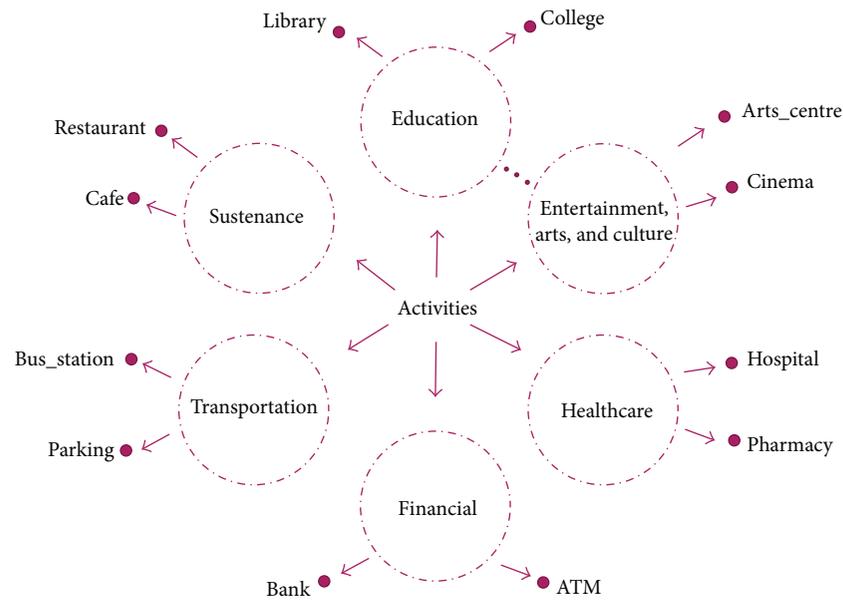
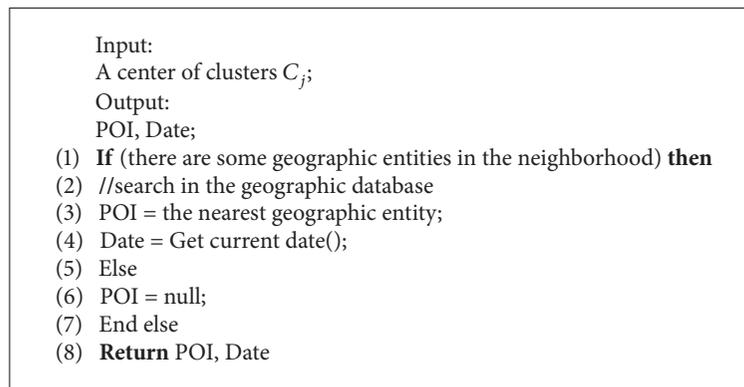


FIGURE 9: The taxonomy of activities.



ALGORITHM 2: Spatial recognition.

For example, the “going to college” activity can be specialized in “education.” Subsequently, for example, if the nearest geographic entity to  $C_i$  has a tag “amenity = library” we deduce that the user is doing an educational activity in the library. This is performed using an algorithm (see Algorithm 3) that searches for the nearest geographic entity and extracts its tags to deduce the activity achieved. This taxonomy is also used to mine quickly users’ past activities; for instance, if we search for the public transportation taken by the user, we just have to specify the activity type as “Transportation” in the SQL request and search for the transportation mode (bus station, taxi station, etc.).

Depending on the application requirements, the activity discovery process can extract further information about the visited POI, such as the address, the number of floors, the opening hours, and the building height. This semantic information permits enhancing the way users perceive and interact with their surroundings; for instance, authors in [30] proposed a system for a mobile navigation assistant fit for the purpose of moving within a complex built space. Their

process of semantic-enhanced POI discovery is based on the exploitation of OSM tags, the same as our method; the results are used in an augmented indoor/outdoor navigation system.

## 4. Experimental Evaluation

In order to test our approach, we will divide the experimentation section into two parts: first, we will test the accuracy of our approach using the Family Coordination dataset [8] by comparing it with CB-SMOT method; then, we will test our approach’s ability to save battery life by comparing our solution to LifeMap application described in [31].

### 4.1. Family Coordination Test

**4.1.1. Dataset.** We used the dataset described in [8] where researchers have conducted experiments in order to reveal the underlying causes of coordination breakdowns that a

```

Input:
POI;
Output:
Activity;
(1) If (POI = null) then
(2) Activity = "no activity";
(3) Else
(4) Tag = Get the tag of POI;
(5) Activity = Search for activity taxonomy (Tag);
(6) End else
(7) Return Activity

```

ALGORITHM 3: Activity discovery.



FIGURE 10: An example of our Spatialite database that contains the geographic entities of Pittsburgh.

routine learning system might be able to address. They tracked GPS locations of the members of six families during six months in Pittsburgh City in Pennsylvania, USA (see Figure 10); researchers made an effort to recruit a wide cross section, selecting families from a variety of ethnic and economic backgrounds, as well as expressing a variety of planning styles, child-rearing models, and transportation preferences. The GPS sampling for every family member including children (a total of 26 persons) was set at one-minute intervals which led to gathering more than two billion and half of GPS points. Moreover, every night during the study, a member of the research team would call the families and interview each parent about that day's management of their kids' activities. In preparation for the interviews, family members were asked to input their daily activities into a web-based survey. Researchers then used the survey to scaffold the phone interview, probing and documenting the overall family logistical plan at each point throughout the day. We

used this information to compare the inferred activities to the real executed activities.

**4.1.2. Test Scenario.** We compared our solution to CB-SMOT algorithm [10] described in the related works; CB-SMOT is widely used to extract visited places from locations' trajectories; we have implemented this algorithm using Weka-STPM platform [32], an extension that adds trajectory processing tools to Weka (Waikato Environment for Knowledge Analysis) and includes CB-SMOT as a clustering algorithm; however, since Weka-STPM is a desktop solution, we have created a mobile version using the CB-SMOT class included in Weka-STPM. The input was the Family Coordination dataset parsed into a Postgresql/PostGIS database and the output was the inferred semantic trajectories (tracking visited places retrieved using Weka-STPM).

Our solution was implemented using Huawei P7 android phone where we have deployed our application that contains the online recognition algorithm, Family Coordination dataset, and a SQLite database that contains the geographic entities of Pittsburgh city needed in the spatial recognition process. Note that, as said previously, this first part of the experimentation is dedicated to testing our approach in terms of accuracy; consequently we will suppose that the phone's battery is fully charged during all the processes principally because we do not have access to the battery's life data in Family Coordination dataset. The ability to save the phone's battery and its impact on the accuracy will be presented in the second part.

**4.1.3. Results.** Results presented in Table 4 represent the comparison of our approach with three versions of CB-SMOT algorithm, by varying its MinTime parameter each time; it was set to 60 s in version 1, to 180 s in version 2, and to 500 s in version 3.

We have tested 10525 activities gathered from the activities of the 24 members of families.

Correct activities represent the number of activities recognized successfully, missed activities represent the number of activities that the users did but the algorithms have failed to recognize, and false activities represent the number of meaningless discovered activities like recognizing the stop of a car in a traffic jam as going to gas stations. The accuracy and the error are calculated as follows:

$$\begin{aligned}
 \text{Accuracy} &= \frac{\text{Correct}}{\text{number of tested activities}}, \\
 \text{Error} &= \frac{\text{missed} + \text{false}}{\text{number of tested activities}}.
 \end{aligned} \tag{8}$$

Globally, our algorithm behaves better than CB-SMOT in terms of accuracy and error. We have recorded the accuracy of 78% in our solution and 61% for the average accuracy of the three versions of CB-SMOT. On the other hand, our error was around 22%, while the average error of the three versions of

TABLE 4: Comparison of our approach to CB-SMOT algorithm.

	Our approach	CB-SMOT V1 MinTime = 90 s	CB-SMOT V2 MinTime = 180 s	CB-SMOT V2 MinTime = 500 s
Tested activities	10525	10525	10525	10525
Correct	8313	4028	7157	8257
Missed	1812	6497	3368	2268
False	578	181	909	2789
Accuracy	78%	38%	68%	78%
Error	22%	63%	40%	48%



FIGURE 11: An example of one user's GPS locations projected on the geographic entities layer of Pittsburgh; (A) stationary activity, (B) moving activity, and (C) moves.

CB-SMOT was around 50%; these results are explained in the following points:

- (i) Our algorithm succeeds in recognizing more activities because it supports the recognition of moving activities; contrariwise CB-SMOT recognizes only stop and move activities (see Figure 11).
- (ii) CB-SMOT shows a higher error because of the MinTime threshold that is set manually (the time of staying from which the algorithm considers that the user has visited this place). Setting this threshold to a small value will increase the number of false alarms (see CB-SMOT V1 in Table 4), while setting it to high value will increase the number of missed activities (see CB-SMOT V2 in Table 4), and, in both cases, the error will be increased since it is calculated on the basis of the number of false and missed activities.

After having tested the accuracy of our approach, we will go further in the experimentation of our battery saving technique.

## 4.2. LifeMap Test

**4.2.1. Dataset.** Researchers in LifeMap project collected real traces from 68 persons over four weeks using HTC Hero, HTC Desire, and Samsung Galaxy S smartphones. The tracking application (called LifeMap) was running as a background service to automatically collect the user's mobility and to trace sensor usage time. To collect the ground truth, the participants explicitly labeled the place names and kept a diary of places they had visited with the entrance and departure times. Moreover, the advantage of using such dataset is the ability to compare the power consumption of our method to the authors' one, since authors tracked the battery status during all the experimentation process.

**4.2.2. Test Scenario.** In this step we will compare our approach to the LifeMap application used to recognize users' motility; the project can be found in [2], the LifeMap dataset can be found in [33], and the LifeMap mobile application can be found on android play store.

We used the LifeMap dataset to test our battery-aware approach; to do so, we developed an android application that is fed by LifeMap datasets; the main idea is to make it out as if the users have moved holding our application in their phones; the application recuperates the GPS coordinates one by one and processes each point using our online approach; after that we compare the battery consumption of our application with LifeMap application's one (see Figure 12).

LifeMap application uses a set of sensors to recognize users' activities; these sensors include GPS, accelerometer, digital compass, and Wi-Fi; the application uses a combination of these sensors to retrieve the activity performed by the user; it is clear that the use of all these sensors represents a major source of power consumption. Consequently, in order to bring an objective comparison between the LifeMap application and our application, it is logical that we have to include the same sensors even if we use only the GPS sensor. Indeed, we have enriched our application with the same sensors used in LifeMap application but the processing of this information is made under the rules defined by our battery-aware approach; for example, the processing of data is made on the estimated  $T_{\min}$  that is calculated on the basis of the remaining battery level and the user's behavior.

Moreover, even if the GPS is activated, the application consumes more resources when it requests the users' position

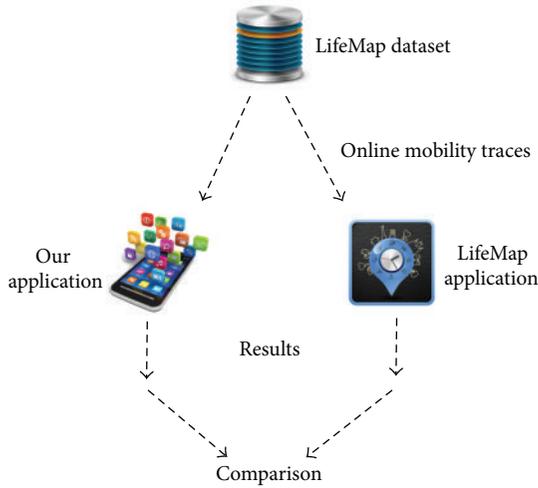


FIGURE 12: Test scenario to compare our solution with LifeMap application using LifeMap dataset.

from the GPS satellites (the case of LifeMap) than when it retrieves it from a database (the case of our application); consequently, in order to compare objectively the two approaches, we continue to obtain the user position from the database; however, on every point retrieved, we request the GPS satellite for a GPS position. Surely, this position is useless but it enables consuming the same resources used by LifeMap application to get a GPS position.

In order to automatically recognize users' activities, we have constructed a Spatialite geodatabase [34] where we stored the background geographic data needed in our spatial recognition (see Figure 13). Note that users' locations are recorded on different frequencies; authors in LifeMap dataset have linked the activity declared by the user to the sampling frequency in order to minimize the size of the database and the power consumption, which has sometimes led to low frequencies (e.g., every 10 minutes). In our approach, we try to recognize accurately users' activities while protecting their phones' battery; to test this service, we need to increase the sampling frequency to see if our work well behaves using a lot of location points. Consequently, we have implemented an interpolation algorithm that estimates the missing locations when the sampling frequency is high (more than one minute).

Our application was deployed on Samsung Galaxy S smartphone, one of the smartphones' models used by LifeMap application to recognize activity.

After having filled the necessities regarding the comparison between the two approaches, we will present the results of both of them.

**4.2.3. Test Results.** The total number of recorded hours of battery status in LifeMap dataset is 48900 hours; in order to experiment on all the hours using one smartphone, we will need over five years of experimentations. Accordingly, we chose to use one random day of each user (rather than 30 days) using five smartphones; the total number of hours compared is 1632 hours. Due to insufficient space, we present



FIGURE 13: An example of one user's trajectory projected on (1) OpenStreetMap raster layer and (2) vector layer retrieved from our geodatabase.

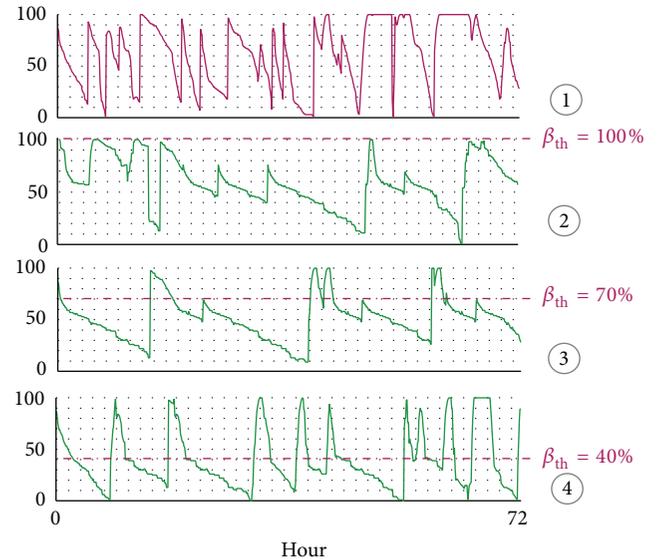


FIGURE 14: Results comparison between LifeMap and our solution for 72 hours of activity recognition. (1) LifeMap, (2) our solution using  $\beta_{th} = 100\%$ , (3) our solution using  $\beta_{th} = 70\%$ , and (4) our solution using  $\beta_{th} = 40\%$ .

in Figure 14 the tracking of one user's battery life for 72 hours using LifeMap and three versions of our approach where we vary each time the value of the threshold  $\beta_{th}$  from where the application starts to save battery life.

TABLE 5: The impact of varying  $\beta_{th}$  on the accuracy.

	$\beta_{th} = 100\%$	$\beta_{th} = 70\%$	$\beta_{th} = 40\%$	LifeMap
Accuracy	68.7%	77.1%	85.4%	78%

Our approach shows an interesting battery saving capacity. Globally, it consumes less resources than LifeMap even when  $\beta_{th}$  is set to a low level (40%).

We notice that, in our approach, battery consumption varies on the basis of  $\beta_{th}$ . When it is set to 100%, our algorithm starts to save the battery from the first moments, which explains the long battery life noticed in graph (2) of Figure 14. However, when  $\beta_{th}$  is set to 40%, we notice that the life of the battery (graph (4)) looks a bit like LifeMap's one (graph (1)); in this case, the algorithm consumption behaves like LifeMap between 100% and 40%, but when it falls under 40%, the algorithm starts saving batteries (see the horizontal dotted line in graph (4)).

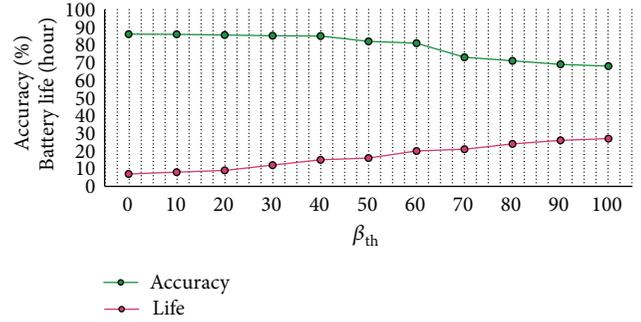
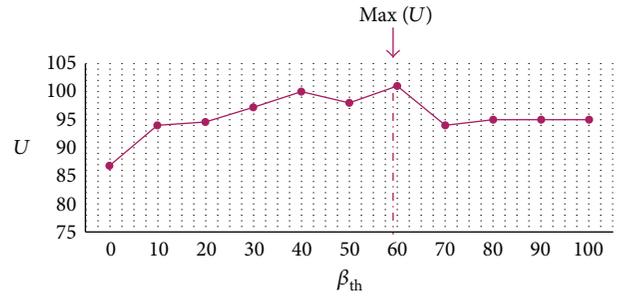
Thus,  $\beta_{th}$  has an impact on the resource consumption, but what about the accuracy? Is it affected by the change of  $\beta_{th}$ ? To answer these questions, we have tracked the accuracy of our algorithm when varying  $\beta_{th}$ ; results are presented in Table 5.

The authors of LifeMap reported in [35] that the accuracy is around 78%. In our case, when  $\beta_{th}$  is set to a value less than 70%, the accuracy of our approach exceeds the LifeMap's one (see Table 5); this is justified by the fact that the LifeMap's technique for detecting important places is based on the time spent by the user around the POI. When a user stays at a given location for more than 10 minutes, the user state is considered stationary and the place is labeled as a POI. This technique fails to recognize moving activities that require a movement to be executed; furthermore, using a fixed time threshold leads to missing some short activities or falsely detecting some activities when the time threshold is set to a small value.

Varying  $\beta_{th}$  has an impact on the accuracy; the more the  $\beta_{th}$  is set to high value, the more the accuracy drops; this is justified by the fact that when  $\beta_{th}$  takes a high value, the temporal window TW shrinks to take less examples and the next processing time  $T_{min}$  will be set further. So if we try to save the battery, we will automatically lose a little bit of accuracy, but is there a compromise between  $\beta_{th}$  and the accuracy? Is it possible to find a  $\beta_{th}$  value that saves as much as possible the accuracy and the battery at the same time?

We have tracked the accuracy of one user's data from LifeMap dataset when varying slowly  $\beta_{th}$ ; results are presented in Figure 15 where the accuracy and an estimation of battery life are presented on the basis of  $\beta_{th}$ . We notice that the higher the value of  $\beta_{th}$  is, the more we save resources and lose precision.

In order to find an optimal solution that enables saving both battery life and accuracy, we have to resolve the optimization problem that maximizes the battery life function  $L(\beta_{th})$  and the accuracy function  $A(\beta_{th})$  at the same time; this is a multiobjective optimization problem where we try to optimize two different functions; one of the existing solutions is the weighted sum method [36]; the solution is based on

FIGURE 15: The impact of varying  $\beta_{th}$  on the accuracy.FIGURE 16: Optimization of  $\beta_{th}$  to maximize the accuracy and the battery life.

selecting a scalar weight  $s_i$  and maximizing the following composite objective function:

$$U = s_1 * L(\beta_{th}) + s_2 * A(\beta_{th}). \quad (9)$$

In our case the importance of the accuracy and the battery life are equal, so the weights of the two variables  $s_1$  and  $s_2$  are equal too ( $s_1 = s_2 = 1$ ); after addition, the value of  $\beta_{th}$  that maximizes  $U$  is  $\beta_{th} = 60\%$  (see Figure 16). Thus, if we need to maximize the accuracy and the battery life at the same time,  $\beta_{th}$  should be set around 60%.

The average precision of our approach when varying  $\beta_{th}$  from 0% to 100% is 79% (see Figure 15); for a system that proposes an activity recognition system keeping a long lifetime of the battery, we think that these results are promising.

The threshold  $\beta_{th}$  from which the system starts to save the phone's battery can be set automatically depending on the activity performed. Indeed, after having learned users' habits, we can link  $\beta_{th}$  to the predicted activity; for instance, when we predict that the user is going to spend a small time outdoors, we set  $\beta_{th}$  to promote the accuracy and vice versa.

## 5. Conclusions and Future Works

In this paper we proposed a new battery-aware technique for extracting semantically and incrementally important geographical locations from users' moves. We associate the places visited by individuals during their movements to meaningful human activities using an algorithm that clusters incrementally user's moves into different types of activities using two parameters, speed and the variance of the moves

orientation. After detecting an important accumulation in a cluster's weight, the cluster is summarized into one point and another process will be launched that aims to search the most meaningful geographic entity near this point (POI); when found, we associate a semantic activity to it. Aiming to save phone batteries, our algorithm was implemented to change its computational complexity on the basis of the remaining battery level and the users' behaviors.

Our approach has been experimented in a real case study to test the accuracy of our recognition mechanism and to observe the impact of our technique on phone's resources. These tests demonstrate that, with a minimum of information, our proposals are capable of online recognizing a person's activities without depleting the phone resources.

Several promising directions for future works exist. First is the enhancement of spatial recognition process with the introduction of probability to assign a cluster to a geographic entity; this probability approach can take advantage of previously recognized activities; for example, a person doing tourist activities all day has more probability to finish his day in a restaurant or in a hotel than in other places. The second enhancement that can be applied to our inferring process is users' profiling. In fact move's pattern of individuals varies between young and old, healthy and sick, and male and female. The implementation of such reflections will improve the accuracy of our activity recognition process.

## Competing Interests

The authors declare that there are no competing interests regarding the publication of this paper.

## References

- [1] J. H. Kang, W. Welbourne, B. Stewart, and G. Borriello, "Extracting places from traces of locations," in *Proceedings of the 2nd ACM International Workshop on Wireless Mobile Applications and Services on WLAN Hotspot (WMASH '04)*, pp. 110–118, Philadelphia, Pa, USA, September 2004.
- [2] *LifeMap: Mobility Monitoring Tool*, 2016, <http://sourceforge.net/projects/lifemapandroid/>.
- [3] A. T. Campbell, S. B. Eisenman, K. Fodor et al., "Transforming the social networking experience with sensing presence from mobile phones," in *Proceedings of the 6th ACM Conference on Embedded Networked Sensor Systems (SenSys '08)*, pp. 367–368, ACM, New York, NY, USA, November 2008.
- [4] N. J. Yuan, Y. Zheng, X. Xie, Y. Wang, K. Zheng, and H. Xiong, "Discovering urban functional zones using latent activity trajectories," *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 3, pp. 712–725, 2015.
- [5] N. J. Yuan, Y. Zheng, L. Zhang, and X. Xie, "T-finder: a recommender system for finding passengers and vacant taxis," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 10, pp. 2390–2403, 2013.
- [6] Y. Zheng, "Trajectory data mining: an overview," *ACM Transactions on Intelligent Systems and Technology*, vol. 6, no. 3, article 29, 2015.
- [7] Y. Li, Y. Zheng, H. Zhang, and L. Chen, "Traffic prediction in a bike-sharing system," in *Proceedings of the 23rd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (GIS '15)*, 2015.
- [8] S. Davidoff, J. Zimmerman, and A. K. Dey, "How routine learners can support family coordination," in *Proceedings of the 28th Annual CHI Conference on Human Factors in Computing Systems (CHI '10)*, pp. 2461–2470, Atlanta, Ga, USA, April 2010.
- [9] N. D. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, and A. T. Campbell, "A survey of mobile phone sensing," *IEEE Communications Magazine*, vol. 48, no. 9, pp. 140–150, 2010.
- [10] A. Tietbohl, V. Bogorny, B. Kuijpers, and L. O. Alvares, "A clustering-based approach for discovering interesting places in trajectories," in *Proceedings of the 23rd Annual ACM Symposium on Applied Computing (SAC '08)*, pp. 863–868, ACM Press, Ceara, Brazil, March 2008.
- [11] Y. Takeuchi and M. Sugimoto, "CityVoyager: an outdoor recommendation system based on user location history," in *Ubiquitous Intelligence and Computing: Third International Conference, UIC 2006, Wuhan, China, September 3–6, 2006. Proceedings*, vol. 4159 of *Lecture Notes in Computer Science*, pp. 625–636, Springer, Berlin, Germany, 2006.
- [12] L. Spinsanti, F. Celli, and C. Renso, "Where you stop is who you are: understanding peoples activities by places visited," in *Proceedings of the Behaviour Monitoring and Interpretation Workshop (BMI '10)*, 2010.
- [13] M. A. Viredaz, L. S. Brakmo, and W. R. Hamburger, "Energy management on handheld devices," *ACM Queue*, vol. 1, no. 7, pp. 44–52, 2003.
- [14] J. P. Pillai and K. G. Shin, "Real-time dynamic voltage scaling for low-power embedded operating systems," *ACM SIGOPS Operating Systems Review*, vol. 35, no. 5, pp. 89–102, 2001.
- [15] B. Zhao, H. Aydin, and D. Zhu, "Reliability-aware dynamic voltage scaling for energy-constrained real-time embedded systems," in *Proceedings of the 26th IEEE International Conference on Computer Design (ICCD '08)*, pp. 633–639, Lake Tahoe, Calif, USA, October 2008.
- [16] E. Shih, P. Bahl, and M. J. Sinclair, "Wake on wireless: an event driven energy saving strategy for battery operated devices," in *Proceedings of the 8th Annual International Conference on Mobile Computing and Networking (MobiCom '02)*, pp. 160–171, Atlanta, Ga, USA, September 2002.
- [17] Y. Wang, J. Lin, M. Annavaram et al., "A framework of energy efficient mobile sensing for automatic user state recognition," in *Proceedings of the 7th International Conference on Mobile Systems, Applications, and Services (MobiSys '09)*, pp. 179–192, ACM, June 2009.
- [18] F. Cao, M. Ester, W. Qian, and A. Zhou, "Density-based clustering over an evolving data stream with noise," in *Proceedings of the 6th SIAM International Conference on Data Mining*, pp. 328–339, Bethesda, Md, USA, April 2006.
- [19] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, 1948.
- [20] V. Novak, I. Perfilieva, and J. Mockoř, *Mathematical principles of fuzzy logic*, Kluwer Academic, Dordrecht, The Netherlands, 1999.
- [21] H. Reichenbach, *Elements of Symbolic Logic*, Macmillan, New York, NY, USA, 1947.
- [22] GIS2GO, Mobile offline GIS app, 2016, <http://www.gis2go.com/home-en.html>.
- [23] GISCloud, Mobile offline GIS app, 2016, <http://www.giscloud.com/blog/offline-maps-are-out/>.

- [24] Esri, Company's History, <http://www.esri.com/about-esri#who-we-are>.
- [25] ArcGIS Runtime SDKs, 2016, <https://developers.arcgis.com/arcgis-runtime/>.
- [26] OpenStreetMap, Planet OSM, 2016, <http://wiki.openstreetmap.org/wiki/Planet.osm>.
- [27] OpenStreetMap extracts, New York region, 2016, <http://download.geofabrik.de/north-america/us/new-york.html>.
- [28] D. O'Sullivan and D. J. Unwin, *Geographic Information Analysis*, John Wiley & Sons, Hoboken, NJ, USA, 2003.
- [29] OpenStreetMap, 2016, <http://wiki.openstreetmap.org/>.
- [30] M. Ruta, F. Scioscia, S. Ieva, G. Loseto, and E. Di Sciascio, "Semantic annotation of OpenStreetMap points of interest for mobile discovery and navigation," in *Proceedings of the IEEE 1st International Conference on Mobile Services (MS '12)*, pp. 33–39, Honolulu, Hawaii, USA, June 2012.
- [31] J. Chon and H. Cha, "LifeMap: a smartphone-based context provider for location-based services," *IEEE Pervasive Computing*, vol. 10, no. 2, pp. 58–67, 2011.
- [32] L. O. Alvares, A. Palma, G. Oliveira, and V. Bogorny, "Weka-STPM: from trajectory samples to semantic trajectories," in *Proceedings of the 11th Workshop de Software Livre (WSL '10)*, pp. 164–169, Porto Alegre, Brazil, 2010.
- [33] *LifeMap Dataset*, Yonsei University, 2016, <http://www.crawdad.org/yonsei/lifemap/20120103/>.
- [34] Spatialite Tools, <https://www.gaia-gis.it/fossil/spatialite-tools/wiki?name=OSM+tools>.
- [35] Y. Chon, E. Talipov, H. Shin, and H. Cha, "Mobility prediction-based smartphone energy optimization for everyday location monitoring," in *Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems (SenSys '11)*, pp. 82–95, ACM, Seattle, Wash, USA, November 2011.
- [36] J. Koski, "Multicriteria truss optimization," in *Multicriteria Optimization in Engineering and in the Sciences*, Stadler, Ed., vol. 37 of *Mathematical Concepts and Methods in Science and Engineering*, pp. 263–307, Plenum Press, New York, NY, USA, 1988.

## Research Article

# Characteristics of Relocated Quiet Zones Using Virtual Microphone Algorithm in an Active Headrest System

**Seokhoon Ryu and Young-Sup Lee**

*Department of Embedded Systems Engineering, Incheon National University, 119 Academy-ro, Yeonsu-gu, Incheon 406-772, Republic of Korea*

Correspondence should be addressed to Young-Sup Lee; [yssl@inu.ac.kr](mailto:yssl@inu.ac.kr)

Received 30 January 2016; Accepted 30 March 2016

Academic Editor: Yurong Qian

Copyright © 2016 S. Ryu and Y.-S. Lee. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study displays theoretical and experimental investigation on the characteristics of the relocated zone of quiet by a virtual microphone (VM) based filtered-x LMS (FxLMS) algorithm which can be embedded in a real-time digital controller for an active headrest system. The attenuation changes at the relocated zones of quiet by the variation of the distance between the ear and the error microphone are mainly examined. An active headrest system was implemented for the control experiment at a chair and consists of two (left and right) secondary loudspeakers, two error microphones, two observer microphones at ear positions in a HATS, and other electronics including a dSPACE 1401 controller. The VM based FxLMS algorithm achieved an attenuation of about 22 dB in the control experiment against a narrowband primary noise by the variation of the distance between the ear and the error microphone. The important factors for the algorithm are discussed as well.

## 1. Introduction

With the advancement of the embedded processors, active noise control (ANC) to cancel an unwanted noise with antinoise by taking the principle of superposition has been studied in many industrial applications [1–5]. In particular, in the case of a vehicle engine, the adaptive feedforward algorithms such as the FxLMS (filtered-reference least mean square) have been applied successfully to actively control a narrowband interior noise in a car cabin generated by the engine [2, 3, 5]. In terms of noise suppression performance, although the global reduction of the engine noise in a car cabin is ideal, this leads to excessive increase of cost, weight, occupancy, and so forth. For this reason, it can be even more efficient to develop local quiet zones [2, 5, 6] near the headrest of chairs. Although the local quiet zone is successfully generated at the location, the head movement in the headrest is restricted due to error microphones near the secondary loudspeaker under the FxLMS approach.

The virtual microphone (VM) technique that relocates the zone of quiet generated at an error microphone to the position of a virtual microphone is introduced to minimize the signal at the virtual microphone [6–9]. Elliott and David

suggested a virtual microphone method for local ANC [6]; likewise Roure and Albarrazin showed a remote microphone technique for ANC [7]. Pawelczyk investigated the multiple adaptive feedback algorithm for active headrest [8] and Moreau et al. reviewed virtual sensing algorithms in ANC [9]. Also a feedback multichannel minimum-variance virtual microphone active noise control system for active headrest system was investigated [10]. Das et al. carried out a remote microphone algorithm using an internal model without using reference microphone [11]. From these studies, the quiet zone formed at the error microphone can be shifted to the ear position by the VM sensing algorithm.

In this paper theoretical and experimental investigation on the characteristics of the relocated zone of quiet by a VM based FxLMS algorithm which can be embedded in a real-time digital controller for an active headrest system is considered against the narrowband swept primary noise such as accelerating car engine noise. The properties of the attenuations and responses at the relocated zones of quiet with the variation of the distance between the observer microphone at the ear position and the error microphone are analyzed in depth. In addition, the governing factors of the VM based FxLMS, such as the secondary path and the virtual

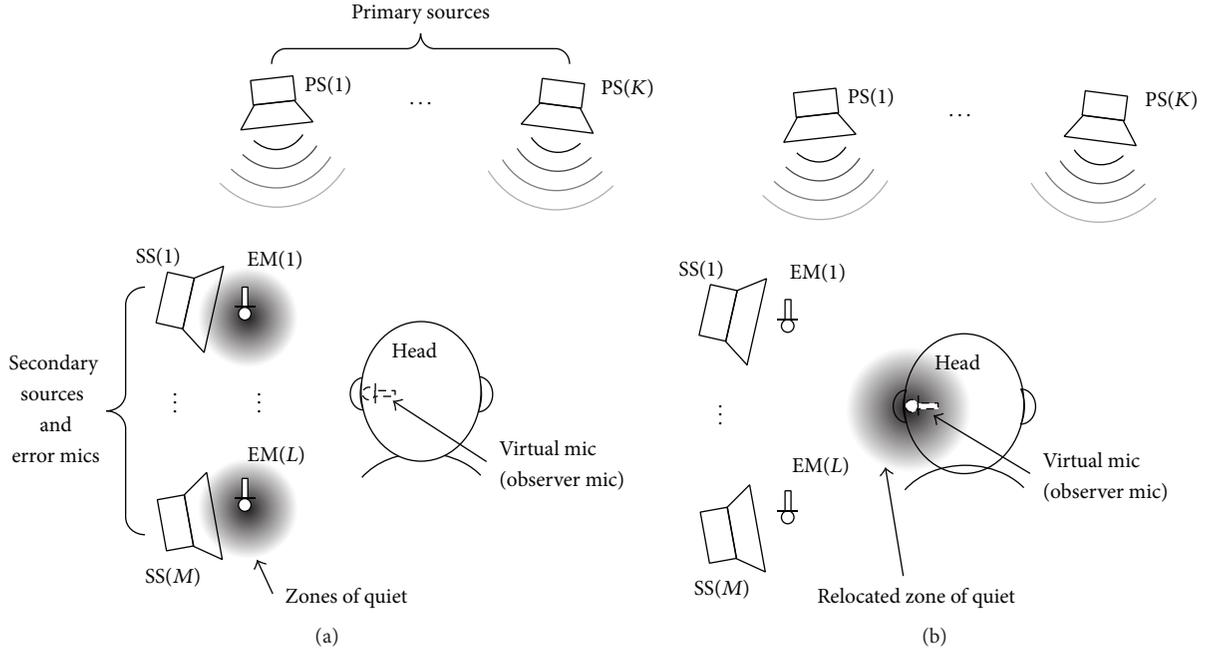


FIGURE 1: Localization of the quiet zone by an active method. (a) Location of the quiet zones at error microphones (VM algorithm off). (b) Relocation of the quiet zone at the ear position (VM algorithm on).

secondary path models, are clarified and their effects are analyzed. The effect of the inclusion of the response difference in the algorithm, which will be symbolized as  $\widehat{V}(z)$  in the main text, between the error microphone and the observer microphone by the primary source is discussed to provide better understanding of the VM based algorithm.

The rest of this paper is organized as follows. In Section 2, theoretical considerations are described to derive proper equations for relocating the quiet zone using the VM based FxLMS algorithm. Section 3 presents the experimental setup including the implementation of the active headrest system, the modelling of the secondary and virtual secondary paths, and the property of the narrowband primary noise. Intensive analysis and discussion from the measured results of the real-time control are written in Section 4. Also, the characteristics of the relocated quiet zones generated by the VM based FxLMS algorithm are discussed in the same section. Finally, conclusions are summarized in Section 5.

## 2. Relocation of Active Quiet Zone Using VM Based FxLMS Algorithm

In this section, the FxLMS algorithm and the VM based FxLMS algorithm are discussed for generating a quiet zone at the position of an error microphone and relocating the quiet zone to the point of a virtual microphone, respectively, in a headrest system.

As illustrated in Figure 1, the active quiet zone around a headrest can be relocated from the error microphone (near to the secondary loudspeaker) to the virtual (observer) microphone (ear location) by using the VM based FxLMS algorithm.

The block diagram of the VM based FxLMS algorithm is shown in Figure 2, where  $P(s)$ ,  $S(s)$ , and  $\widehat{S}(z)$  are the primary path, the secondary path, and the secondary path model, respectively. The subscript  $v$  in  $P_v(s)$ ,  $S_v(s)$ , and  $\widehat{S}_v(z)$  represents the virtual path. Also  $W(z)$  is the digital control filter which can be updated at every iteration with the FxLMS algorithm and  $\widehat{V}(z)$  the digital response difference between the error microphone and the virtual microphone when a secondary source generates white noise. As the secondary path model  $\widehat{S}(z)$  and the virtual secondary path model  $\widehat{S}_v(z)$  are implemented as FIR filters,  $\widehat{V}(z)$  is also designed as an FIR filter. The exactness of the model  $\widehat{V}(z)$  is important in minimizing the error signal for the fast adaptation of the algorithm. The process inside of the dashed line in Figure 2 is the digital part that is worked in a DSP and the outside of the dashed line is the analogue part. The reference signal  $x(t)$  with narrowband property is converted to the form of the discrete reference signal  $x(n)$  after passing through the ADC (analogue-to-digital converter)—frequency estimation—Sine and Cosine (wave generator) as illustrated in Figure 2. The summation of the disturbance  $d(t)$  and the plant output signal  $y(t)$  at the error microphone is the error signal which is given as  $e(t) = d(t) + y(t)$ . Also the virtual error signal at the virtual (observer) microphone can be written by  $e_v(t) = d_v(t) + y_v(t)$ .

As presented in Figure 2, the VM based FxLMS algorithm relocates the zone of quiet to the observer (virtual) microphone from the error microphone. Thus the virtual error signal  $e_v(t)$  at the listener's ear position is necessary to operate the VM based FxLMS algorithm. However it is impossible to obtain the virtual error signal  $e_v(t)$  from the listener's ear in the real-time control of a practical active headrest system; the virtual error signal must be estimated



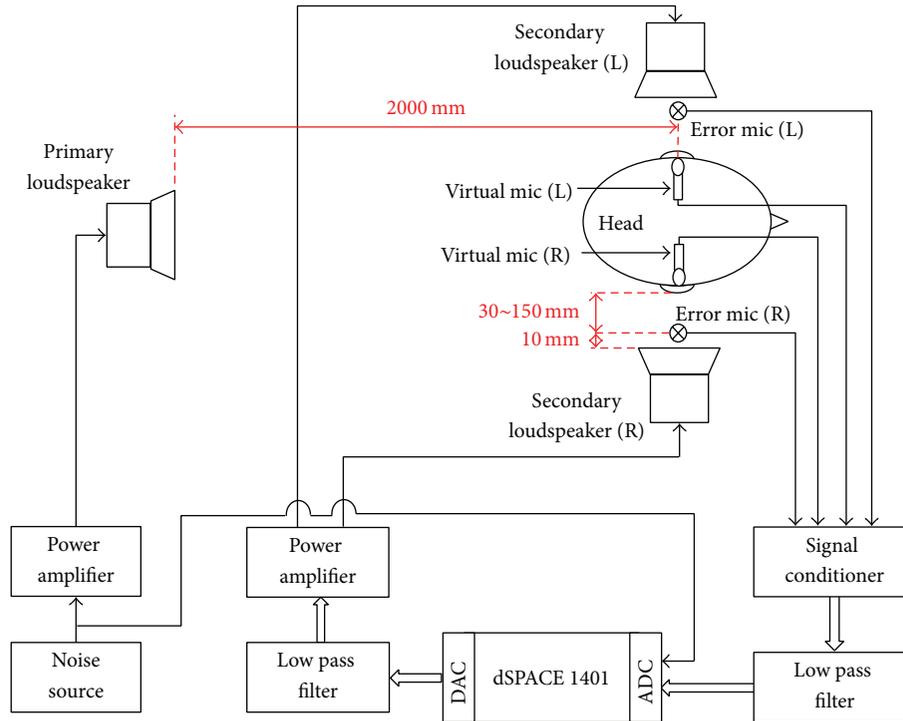


FIGURE 3: Schematic diagram of the active headrest system for control experiment.

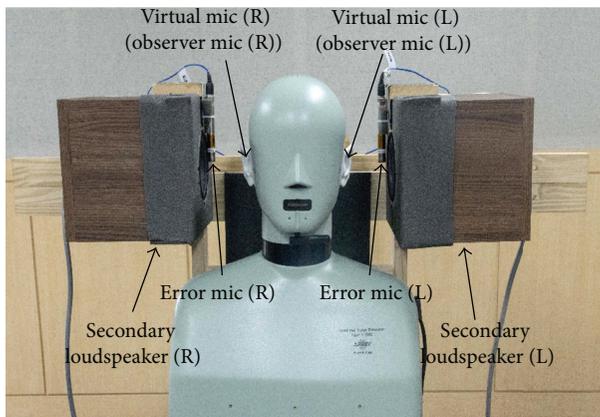


FIGURE 4: Implementation of the active headrest system.

between the primary loudspeaker and the error microphone or the observer microphone was 2000 mm. The distance  $\Delta$  between the observer microphone and the error microphone, which was fixed with the secondary loudspeaker at the distance of 10 mm, was varied from 30 to 150 mm with the interval of 30 mm for the measurement of the quiet zone relocation. Thus the distances were  $\Delta = 30, 60, 90, 120,$  and  $150$  mm.

For the implementation of the real-time VM based FxLMS for the active headrest, a dSPACE 1401 is used as an embedded controller with some other electronics and circuits including the low pass filters with the cutoff frequency of about 2 kHz for antialiasing and reconstruction filtering of

the signals, two signal conditioners (PCB 480E09), and a power amplifier as shown in Figure 3. The sampling frequency was 10 kHz in the real-time control of the headrest system.

The primary noise was a linearly swept narrowband signal with four-component orders of the fundamental frequency C1 (200 → 320 Hz), C1.5 (300 → 480 Hz), C2 (400 → 640 Hz), and C2.5 (500 → 800 Hz). The primary noise which is generated by the powertrain of a passenger car is considered in this study. The noise is a critical problem of such a car and is composed of multiple orders of a sinusoidal wave. The first four orders C1, C1.5, C2, and C2.5 which are the largest contribution of the interior noise are investigated. The noise signal held the lowest frequencies for 3 seconds at the beginning and was swept for 10 seconds at the next step and then maintained the highest frequencies for 2 seconds.

**3.2. Secondary Path Modelling.** As shown in Figure 2, the active headrest system requires the digital models for the secondary path (secondary loudspeaker, near error microphone)  $\hat{S}(z)$  and the virtual secondary path (secondary loudspeaker, near observer microphone)  $\hat{S}_v(z)$ , for the implementation of the VM based FxLMS algorithm in dSPACE 1401. The two secondary paths were implemented as impulse response functions (IRF)  $\hat{s}(n)$  and  $\hat{s}_v(n)$ , respectively. By the way the cross secondary path (secondary loudspeaker, opposite error microphone) and the cross virtual secondary path (secondary loudspeaker, opposite observer microphone) were ignored in the control algorithm since they are very small compared to the secondary path and the virtual secondary path, respectively, as plotted in Figure 5.

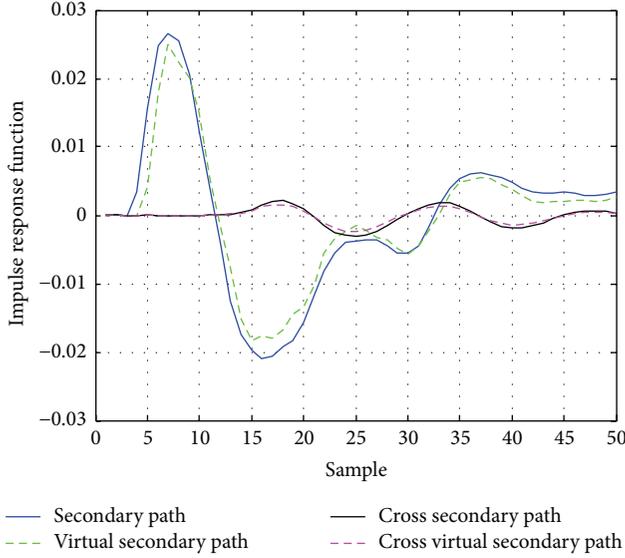


FIGURE 5: Impulse response functions of the secondary, virtual secondary, cross secondary, and cross virtual secondary paths at the distance  $\Delta = 30$  mm.

## 4. Experiment Results and Discussions

**4.1. Performance Comparison of FxLMS and VM Based FxLMS.** In the real-time control experiment, the two ANC algorithms of the FxLMS and the VM based FxLMS were considered for the active headrest system.

In Figures 6(a) and 6(b), the power spectral densities (PSDs) before and after control,  $S_{ee,before}(f)$  and  $S_{ee,after}(f)$ , and the attenuation,  $A(f) = 10 \log_{10}[S_{ee,after}(f)/S_{ee,before}(f)]$ , of the error signal  $e(t)$ , as shown in Figure 2, measured at the error microphone are plotted, respectively, when the FxLMS algorithm is applied in the headrest system. Both plots show the responses after control are similar when the distances  $\Delta$  between the error microphone and the observer (virtual) microphone are 30, 60, 90, 120, and 150 mm. As the frequency range of the narrowband primary noise with the four orders (C1, C1.5, C2, and C2.5) is 200–800 Hz, the plots are dominant at the frequency range. The amounts of the averaged noise attenuation levels after control are about 14–31 dB over the frequency range regardless of the variation of  $\Delta = 30$ –150 mm as shown in Figure 6 since the quiet zones are produced at the error microphone positions by the FxLMS algorithm.

In Figures 7(a) and 7(b), the PSDs,  $S_{e_v,e_v,before}(f)$  and  $S_{e_v,e_v,after}(f)$ , and the attenuation,  $A_v(f) = 10 \log_{10}[S_{e_v,e_v,after}(f)/S_{e_v,e_v,before}(f)]$ , of the observer signal  $e_v(t)$ , as shown in Figure 2, measured at the observer microphone (ear position) are displayed, respectively, when the FxLMS algorithm is implemented. In this case, the noise attenuation levels after control are dramatically decreased by the increase of the distance  $\Delta$ . Since the location of the quiet zone is centered at the error microphone, the increase of the distance  $\Delta$  indicates the observer microphone is away further from the center of the quiet zone as explained in Figure 1. Hence the amounts of the averaged attenuation levels at  $\Delta = 30, 60, 90, 120,$  and  $150$  mm are about 12, 6,

4, 2, and 1 dB, respectively, as illustrated in Figure 7. This implies the increased  $\Delta$  causes less reduction at the observer microphone in control.

In Figures 8(a) and 8(b), the PSDs,  $S_{e_e,before}(f)$  and  $S_{e_e,after}(f)$ , and the attenuation,  $A(f)$ , of the error signal  $e(t)$  measured at the error microphone are plotted, respectively, when the VM based FxLMS algorithm is applied. Both plots show the responses after control are greatly different when the distance  $\Delta$  varies. The noise attenuation levels after control are dramatically decreased or even enhanced by the increase of the distance  $\Delta$ . Since the center of the relocated quiet zone is the observer microphone (ear position), the increase of the distance  $\Delta$  indicates the error microphone is away further from the center of the relocated quiet zone as explained in Figure 1.

In Figures 9(a) and 9(b), the PSDs,  $S_{e_v,e_v,before}(f)$  and  $S_{e_v,e_v,after}(f)$ , and the attenuation,  $A_v(f)$ , of the observer signal  $e_v(t)$  measured at the observer microphone (ear position) are displayed, respectively, when the VM based FxLMS algorithm is implemented. Both plots show the responses after control are quite different when the distance  $\Delta$  varies. The noise attenuation levels after control are gradually decreased by the increase of the distance  $\Delta$ . Although the virtual secondary paths  $\hat{S}_v(z)$  at each distance  $\Delta$  are modelled and implemented in the VM based FxLMS at the corresponding control experiment, the attenuation levels become worse by the increase of  $\Delta$  as shown in Figure 9(b). This is because  $\hat{S}_v(z)$  with the larger  $\Delta$  contains the larger inaccuracy in the modelling of the actual path  $S_v(z)$ . The amounts of the averaged attenuations at the distance  $\Delta = 30, 60, 90, 120,$  and  $150$  mm are about 20, 22, 19, 15, and 8 dB, respectively. This implies the VM control technique provides further reduction at the observer microphone position in control than the simple FxLMS algorithm.

**4.2. Averaged Attenuation Levels around Quiet Zones.** In Figures 10 and 11, the amounts of the averaged attenuation levels,  $A(f)$  and  $A_v(f)$ , of the measured signals by the error microphone and the observer microphone around the quiet zones after control with both the FxLMS and the VM based FxLMS algorithms, respectively, against the distances  $\Delta = 30, 60, 90, 120,$  and  $150$  mm are outlined. The solid and dotted lines denote the averaged attenuation levels at the error and observer microphones, respectively.

As shown in Figure 10, after the FxLMS control, the amounts of the averaged  $A(f)$  measured at the error microphones are about 23–24 dB regardless of the variation of  $\Delta$  and this is also plotted in Figure 6(b) and discussed in Section 4.1. On the other hand, the amounts of the averaged  $A_v(f)$  measured at the observer microphones are steadily worse from 12 dB to 1 dB with the increase of  $\Delta$ . In addition, the differences between the averaged  $A(f)$  and  $A_v(f)$  vary from 11 dB at  $\Delta = 30$  mm to 22 dB  $\Delta = 150$  mm. Hence, in the FxLMS algorithm, the increased  $\Delta$  causes the less averaged  $A_v(f)$  and the increased difference between the averaged  $A(f)$  and  $A_v(f)$ . It is worthy to note that 6 dB of the averaged  $A_v(f)$  is achieved at about the distance  $\Delta = 60$  mm. Assuming the shape of the quiet zone is a sphere, it is meaningful as the  $\Delta$  value can indicate the radius of the quiet

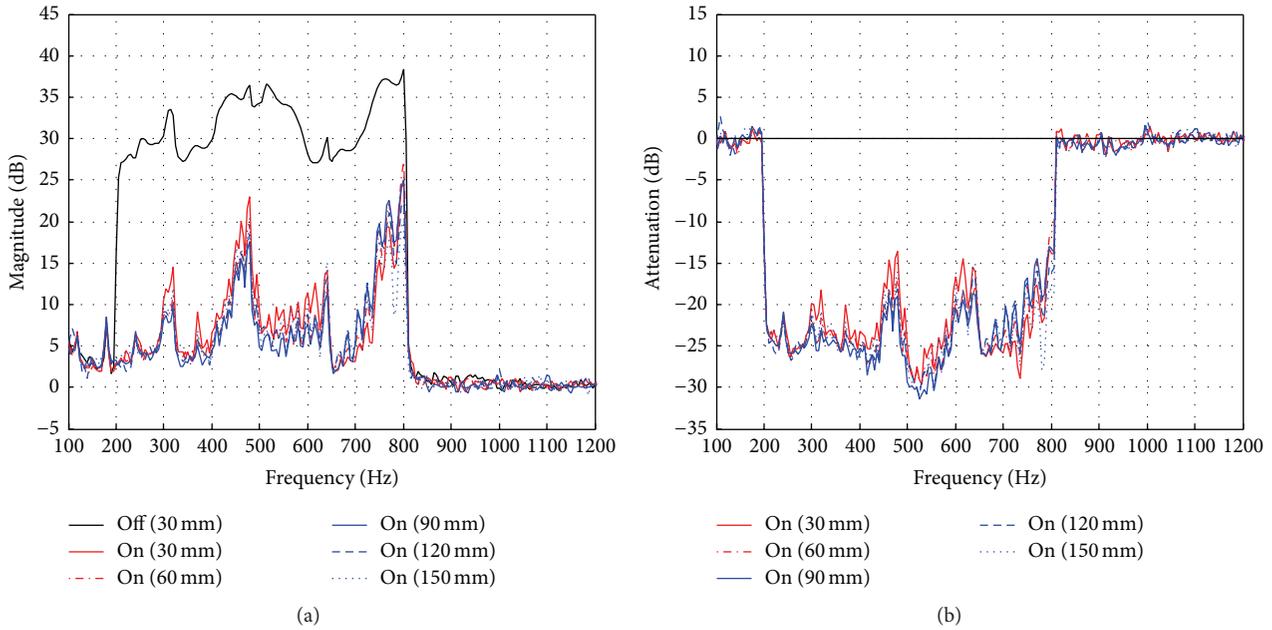


FIGURE 6: Comparison of the measured error signals at the error microphones with the FxLMS algorithm. (a) PSD functions. (b) Attenuations.

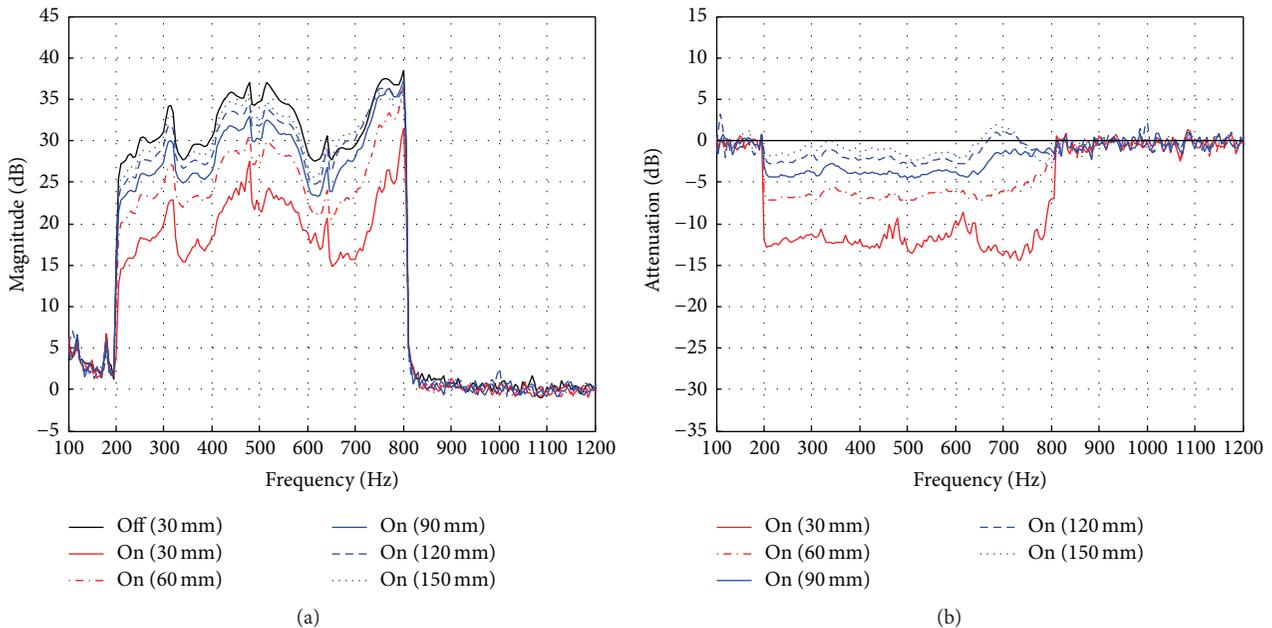


FIGURE 7: Comparison of the measured error signals at the observer (virtual) microphone with the FxLMS algorithm. (a) PSD functions. (b) Attenuations.

zone. The 6 dB attenuation radius of the quiet zone is about 60 mm in this active headrest system against the narrowband primary noise with the FxLMS algorithm.

As displayed in Figure 11, after control with the VM based FxLMS algorithm, the amounts of the averaged  $A_v(f)$  (dotted line) measured at the observer microphones are about 8–22 dB by the variation of  $\Delta$  as the averaged  $A_v(f)$  is gradually worse with the increase of the distance  $\Delta$  although the corresponding virtual secondary path model  $\widehat{S}_v(z)$  at each distance

is applied in the algorithm. Likewise the averaged  $A(f)$  (solid line) measured at the error microphones is steadily worse with the increase of  $\Delta$ . It is notable that the enhancement is measured by the error microphone position with the distance  $\Delta = 80$  mm when the amount of the averaged attenuation at the observer microphone (ear position) is 20 dB. This implies that if a listener at the active headrest system moves his/her ear toward the error microphone under the VM based FxLMS algorithm, he/she will hear even louder noise

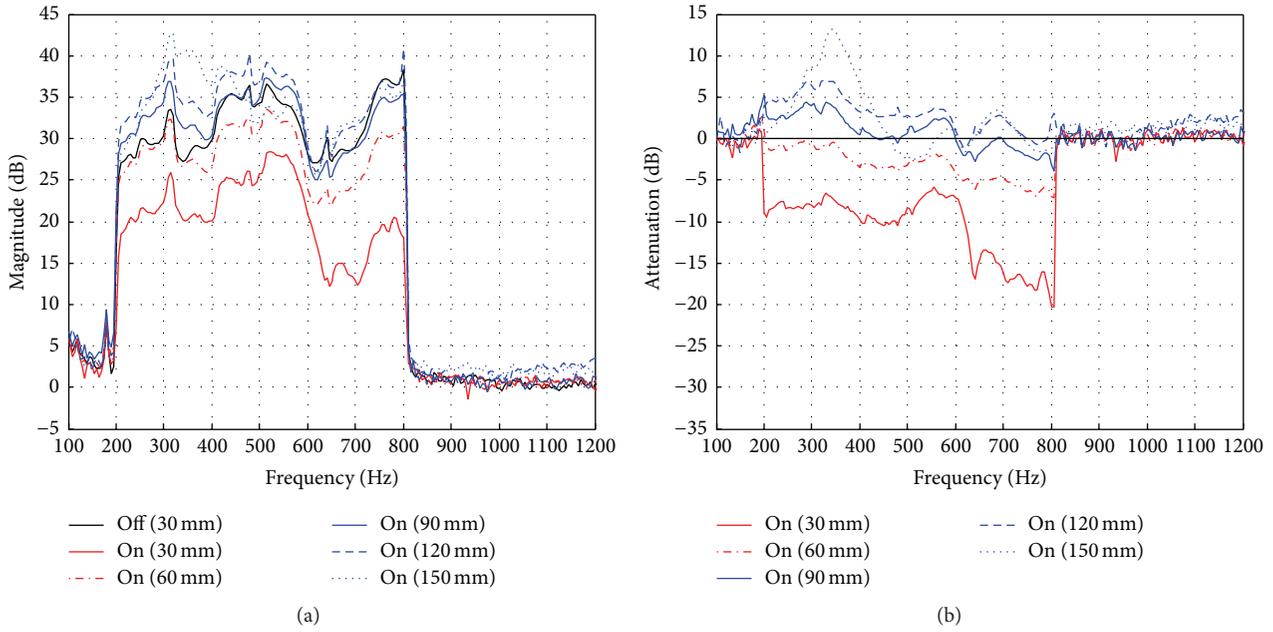


FIGURE 8: Comparison of the measured error signals at the error microphones with the VM based FxLMS algorithm. (a) PSD functions. (b) Attenuations.

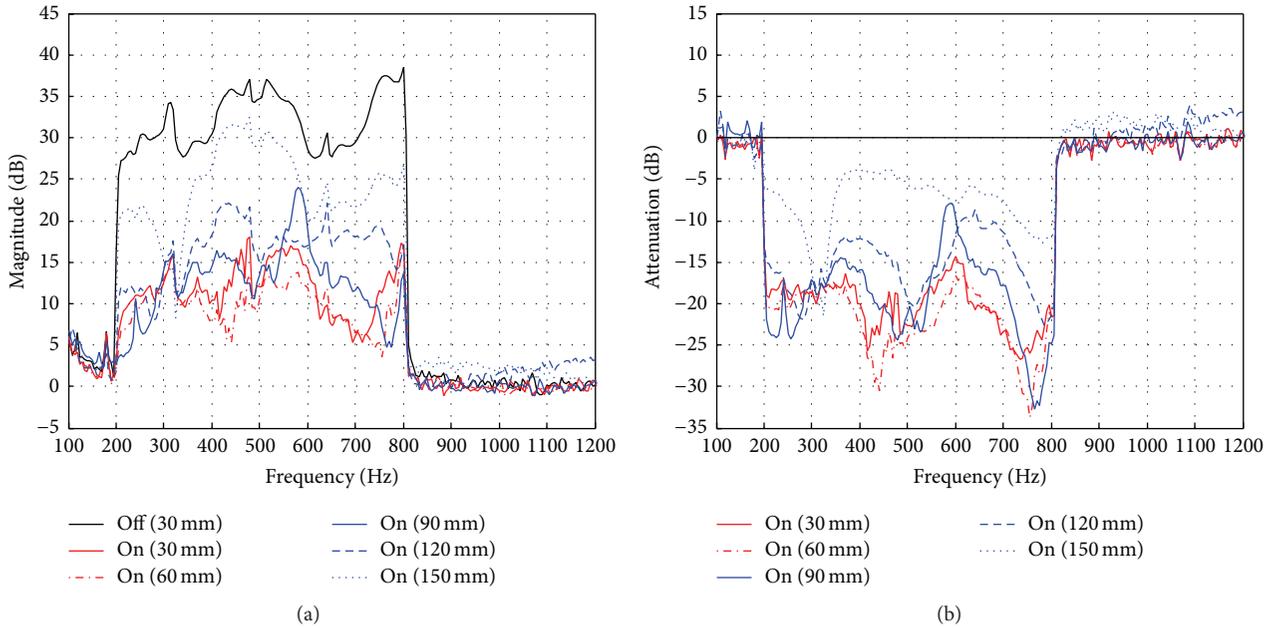


FIGURE 9: Comparison of the measured error signals at the observer (virtual) microphone with the VM based FxLMS algorithm. (a) PSD functions. (b) Attenuations.

compared to the noise before control. Because of this reason, thus, the allowable head movement range of the listener may be restricted in a practical active headrest system.

For the extension of the quiet zones, it is necessary to have even more exact and precise models of the secondary path and the virtual secondary path. Also more error microphones

and secondary loudspeakers can provide further reduction of noise in the headrest system.

In summary of the above discussions, the accuracy of the path models of  $\hat{S}(z)$ ,  $\hat{S}_v(z)$ , and  $\hat{V}(z)$  is important to estimate  $\hat{\mathbf{e}}_v(n)$  in the VM based FxLMS algorithm. Then the algorithm allows achieving reasonable attenuations with proper stability

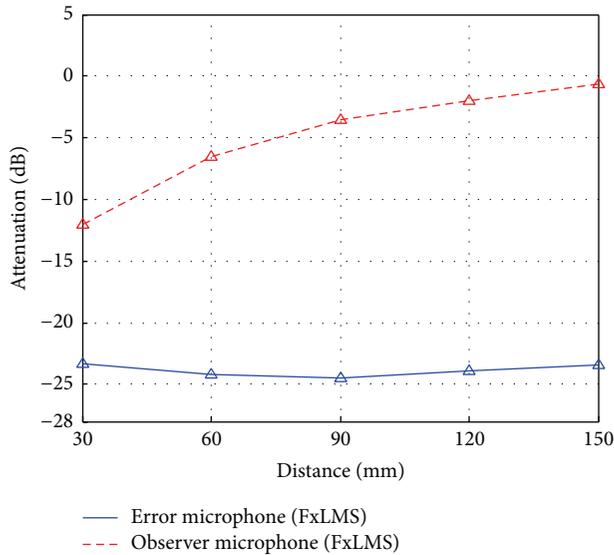


FIGURE 10: Comparison of the averaged attenuations of the measured signals at the error microphone and the observer microphone against the distance  $\Delta$  with the FxLMS algorithm.

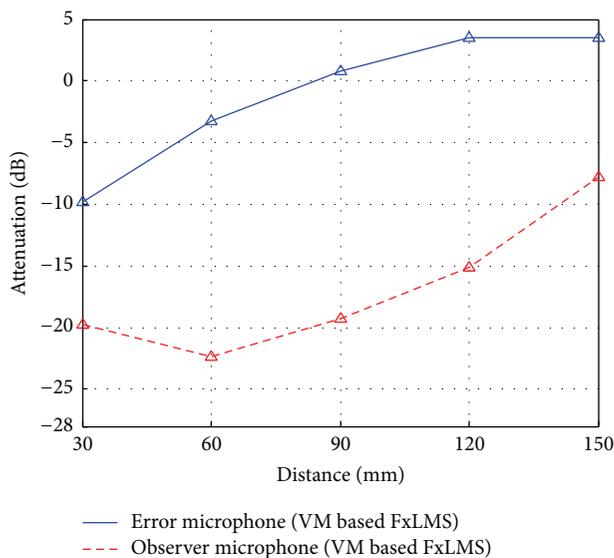


FIGURE 11: Comparison of the averaged attenuations of the measured signals at the error microphone and the observer microphone against the distance  $\Delta$  with the VM based FxLMS algorithm.

at the listener's ear position in the practical applications such as an active headrest system.

## 5. Conclusions

This investigation describes a new VM based FxLMS algorithm which can be embedded in a real-time digital controller for an active headrest system. The main outcomes are summarized in the following.

Different from the FxLMS algorithm, the averaged attenuation  $A_v(f)$  at the ear position (measured by the observer

microphone) achieved was about 8 dB at  $\Delta = 150$  mm and 22 dB at  $\Delta = 60$  mm under the VM based FxLMS algorithm. The amount of  $A_v(f)$  was decreased with the increase of the distance  $\Delta$ .

The suggested VM based FxLMS algorithm requires the path models of  $\widehat{S}(z)$ ,  $\widehat{S}_v(z)$ , and  $\widehat{V}(z)$  and the accuracy of these models is important to estimate  $\widehat{e}_v(n)$ . Then the algorithm allows achieving reasonable attenuations with proper stability at the listener's ear position in the headrest system.

Extending the quiet zones in the active headrest will be considered in terms of multiple error microphones and loudspeakers for the application of real headrest systems in the future.

## Competing Interests

The authors declare that they have no competing interests.

## Acknowledgments

This research was supported by a grant (12-TI-C01) from the Advanced Water Management Research Program funded by Ministry of Land, Infrastructure and Transport of Korean government.

## References

- [1] J. Diaz, J. M. Egaña, and J. Viñolas, "A local active noise control system based on a virtual-microphone technique for railway sleeping vehicle applications," *Mechanical Systems and Signal Processing*, vol. 20, no. 8, pp. 2259–2276, 2006.
- [2] S. J. Elliott, *Signal Processing for Active Control*, Academic Press, 2001.
- [3] S. M. Kuo and D. R. Morgan, *Active Noise Control Systems: Algorithms and DSP Implementations*, John Wiley & Sons, New York, NY, USA, 1996.
- [4] J. Kim, Y. Choi, and Y.-S. Lee, "Spectrogram image analysis of error signals for minimizing impulse noise," *Journal of Sensors*, vol. 2016, Article ID 4935694, 9 pages, 2016.
- [5] S. J. Elliott, I. M. Stothers, P. A. Nelson, A. M. McDonald, D. C. Quinn, and T. Saunders, "The active control of engine noise inside cars," in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings (InterNoise '88)*, vol. 1998, no. 3, pp. 987–990, Avignon, France, 1988.
- [6] S. J. Elliott and A. David, "A virtual microphone arrangement for local active sound control," in *Proceedings of the 1st International Conference on Motion and Vibration Control*, pp. 1027–1031, September 1992.
- [7] A. Roure and A. Albarrazin, "The remote microphone technique for active noise control," in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, vol. 1999, no. 5, pp. 1233–1244, Fort Lauderdale, Fla, USA, 1999.
- [8] M. Pawelczyk, "Multiple input-multiple output adaptive feedback control strategies for the active headrest system: design and real-time implementation," *International Journal of Adaptive Control and Signal Processing*, vol. 17, no. 10, pp. 785–800, 2003.
- [9] D. Moreau, B. Cazzolato, A. Zander, and C. Petersen, "A review of virtual sensing algorithms for active noise control," *Algorithms*, vol. 1, no. 2, pp. 69–99, 2008.

- [10] M. Pawelczyk, "Multi-channel virtual-microphone feedback minimum-variance active noise control system," *Mechanics*, vol. 28, no. 1, pp. 12–17, 2009.
- [11] D. P. Das, D. Moreau, and B. Cazzolato, "Performance evaluation of an active headrest using the remote microphone technique," in *Proceedings of the Annual Conference of the Australian Acoustical Society*, Gold Coast, Australia, 2011, Paper Number 69.

## Research Article

# An Efficient Image Enlargement Method for Image Sensors of Mobile in Embedded Systems

Hua Hua,<sup>1</sup> Xiaomin Yang,<sup>1</sup> Binyu Yan,<sup>1</sup> Kai Zhou,<sup>2</sup> and Wei Lu<sup>3</sup>

<sup>1</sup>College of Electronics and Information Engineering, Sichuan University, Chengdu, Sichuan 610064, China

<sup>2</sup>College of Electrical and Engineering Information, Sichuan University, Chengdu, Sichuan 610064, China

<sup>3</sup>School of Software Engineering, Beijing Jiaotong University, Beijing 100044, China

Correspondence should be addressed to Xiaomin Yang; yangxiaomin\_2001@163.com

Received 20 November 2015; Revised 5 March 2016; Accepted 21 March 2016

Academic Editor: Marco Anisetti

Copyright © 2016 Hua Hua et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Main challenges for image enlargement methods in embedded systems come from the requirements of good performance, low computational cost, and low memory usage. This paper proposes an efficient image enlargement method which can meet these requirements in embedded system. Firstly, to improve the performance of enlargement methods, this method extracts different kind of features for different morphologies with different approaches. Then, various dictionaries based on different kind of features are learned, which represent the image in a more efficient manner. Secondly, to accelerate the enlargement speed and reduce the memory usage, this method divides the atoms of each dictionary into several clusters. For each cluster, separate projection matrix is calculated. This method reformulates the problem as a least squares regression. The high-resolution (HR) images can be reconstructed based on a few projection matrixes. Numerous experiment results show that this method has advantages such as being efficient and real-time and having less memory cost. These advantages make this method easy to implement in mobile embedded system.

## 1. Introduction

Over the last few decades, people have widely adopted mobile phones to life. For 2017, the number of mobile phone users will reach almost 5.3 billion. For many mobile phone users, mobile phone is used not only for spoken communication but also as a tool to capture images. Mobile phones offer great benefits to the users by enabling photography and video recording always and everywhere. Unfortunately, many of the images being taken with mobile phone are low in resolution since the low quality image sensor. There are two ways to obtain high-resolution images: (1) replace the mobile phone with a more powerful mobile phone; (2) use some methods to enlarge the images. Most of the mobile phone users prefer to use a method to enlarge the image rather than replacing the mobile phone with a more powerful mobile phone. Many efforts have been devoted to image enlargement methods in the past decade. However, the enlargement methods face three challenges when applied in embedded systems: (1) performance requirement, (2) real-time requirement, and (3) constraint on memory consumption.

Superresolution (SR) is one of the most prospective image enlargement methods. Existing SR methods can be divided into three categories: interpolation-based methods [1, 2], reconstruction-based methods [3–5], and example-based methods [6–10].

The interpolation-based methods [1, 2] apply the correlation of neighboring image pixels to approximate the fundamental HR pixels. These types of methods have lower computation complexities. However, the interpolation does not add any new detailed information into the enlarged image. The quality of the enlarged image is still unsatisfying and it may cause the aliasing to the enlarged LR image. Although the interpolation-based methods run fast and need little memory, the poor performance limits the application of interpolation-based methods for the implementation in embedded system.

Reconstruction-based methods [3–7] require different LR images of the same scene taken from slightly moved viewpoints, and those LR images have different subpixel shifts from each other. This category of methods tries to exploit additional information from a sequence of successive LR images of the same scene to synthesize HR images. Compared

with interpolation-based methods, the reconstruction-based methods obtain better performance with a small desired magnification factor. However, the performance of this kind of methods degrades rapidly when the desired magnification factor becomes large. The reconstruction-based methods need to store the information of all the sequence LR images, which is high memory requirement. Due to the above reasons, reconstruction-based methods are not ready for embedded system.

Single image SR methods such as neighbor embedding-based methods [7, 8], regression-based methods [9, 10], and sparse representation-based methods [11–15] have been explored in recent years. These methods presume that the high-frequency details lost in the LR images can be predicted through learning the cooccurrence relationship between LR training patches and their corresponding HR patches. Recently, sparse representation-based methods have proven to be effective towards solving image superresolution problems. Yang et al. [16] proposed an approach based on sparse representation, with the assumption that the HR and LR images share the same set of sparse coefficients. Therefore, the HR image can be reconstructed by combining the trained HR dictionary and the sparse coefficients of the corresponding LR image. Although the sparse representation-based methods offer a good performance, the optimization of dictionary learning and image reconstruction has a problem of highly intensive computation. Besides, sparse representation-based SR methods reserve memory to store the information of HR dictionary and LR dictionary. The size of dictionary impacts the memory usage. Sparse representation-based SR methods require intensive large memory, especially with increasing size of dictionary. Both the time-complexity and memory usage are key limit factors in the embedded system applications of these methods. Zhan et al. [17] proposed a fast multiclass dictionaries learning method in MRI reconstruction. Timofte et al. [18] constructed a set of mapping relationships between the LR and HR patches using a learned LR-HR dictionary. Anchored Neighborhood Regression method [18] reformulates the problem as a least squares regression, which leads to a vast computational speedup while keeping the same accuracy as previous methods. Anchored Neighborhood Regression method calculates the mapping matrix based on a universal dictionary. However, a large number of different structural patterns exist in an image, whereas one dictionary is not capable of capturing all of the different morphologies. Besides, Anchored Neighborhood Regression method still needs to store separate projection matrix for each dictionary atom which is high memory usage.

The existing sparse representation-based SR methods always suffer from three main problems for embedded system. First, the performance of these methods is limited, since these methods only use one approach to extract the features of the image for presenting the LR image generally. However, the morphologies vary significantly across images. Different patches prefer different features for accurately representing different morphologies. A single feature extraction approach cannot represent the image accurately. Therefore, jointly representing an image with different kind of features is important. Furthermore, time-complexity and memory usage are

key limit factors in the embedded system applications of these methods. The optimization of dictionary learning and image reconstruction leads to highly intensive computation. Sparse representation-based methods need to reserve memory to store the information of HR dictionary and LR dictionary. The size of dictionary impacts the memory usage.

Above all, this study makes the following three main contributions. (1) Jointly representing an image with different types of features is proposed in feature extraction stage. For accurately representing different morphologies, images (or patches) prefer different types of features extracted by different approaches, since one single feature extraction approach cannot accurately capture the essential features of the image. (2) Multiple dictionaries are learned based on different types of features in sparse representation stage, since one dictionary with single type of features is inadequate in capturing all of the different morphologies of the image. To capture the different morphologies of the image more accurately, multifeature dictionaries, which consist of different dictionaries with different features, are learned. (3) To reduce the computational cost and memory usage, we propose an Anchored Cluster Regression method. Anchored Cluster Regression method divides the dictionary atoms into several clusters. Then, the projection matrix for each cluster is calculated. In Anchored Cluster Regression method, each HR patch can be reconstructed by the projection matrix of its corresponding cluster. Anchored Cluster Regression method reformulates the problem as a least squares regression. It only needs to store the projection matrix of each cluster. Anchored Cluster Regression leads to a vast computational speedup and needs less memory.

## 2. Sparse Representation-Based SR Method

Superresolution aims to reconstruct the HR image from the LR image, which can be formulated as follows:

$$Y = HBX, \quad (1)$$

where  $Y \in R^N$  is the observed low-resolution (LR) image.  $X \in R^{NL}$  is its corresponding high-resolution (HR) image of the same scene.  $Y$  is a downsampled and blurred version of  $X$ .  $H$  denotes a downsampling operator and  $B$  is the blur operator.

Let  $\text{Patch}_-y^i$  be LR patch of the LR image with the size  $\sqrt{n} \times \sqrt{n}$  at the location  $i$ ,  $i = 1, 2, \dots, C$ . Then, we have

$$\text{Patch}_-y^i = R_i(Y), \quad (2)$$

where  $R_i(\cdot)$  is an operator that extracts a patch at position  $i$  from the LR image  $Y$ .

Similarly, the corresponding HR patch  $\text{Patch}_-x^i$  is with the size  $\sqrt{nL} \times \sqrt{nL}$  at the location  $i$ ,  $i = 1, 2, \dots, C$ . And we have

$$\text{Patch}_-x^i = R_i(X). \quad (3)$$

With LR patch  $\text{Patch}_-y^i$ ,  $y^i$  is the feature extracted from  $\text{Patch}_-y^i$ . The feature can be expressed as

$$y^i = \text{Features}_L(\text{Patch}_-y^i), \quad (4)$$

where  $\text{Features}_L(\cdot)$  refers to extracting LR feature operator.

Subsequently, the corresponding HR feature  $x^i$  is extracted from HR patch  $\text{Patch}_{x^i}$ :

$$x^i = \text{Features}_{H_H}(\text{Patch}_{x^i}), \quad (5)$$

where  $\text{Features}_{H_H}(\cdot)$  refers to extracting HR feature operator, which is usually the differences between the LR image and its corresponding HR image.

With the sparse generating model, each LR patch feature  $y^i$  ( $y^i \in \mathfrak{R}^n$ ) can be projected over the LR dictionary  $D_l$ , which characterizes the LR patches. This projection produces a sparse representation of  $y^i$  via  $\alpha_l^i$ :

$$y^i = D_l \alpha_l^i \quad y^i \in \mathfrak{R}^n, \quad \|\alpha_l^i\|_0 \ll k, \quad (6)$$

where  $D_l$  and  $\alpha_l^i$  are the LR dictionary and the sparse representation of  $y^i$ , respectively. Generally, in order to obtain an optimal  $\alpha_l^i$  that has the fewest nonzero elements, we should solve the following optimization problem:

$$\alpha_l^i = \min_{\alpha} \quad \|D_l \alpha - y^i\|_2 + \lambda \|\alpha\|_0, \quad (7)$$

where  $\lambda$  is a constant.

Similarly, we have the sparse representation of the HR patch:

$$x^i \approx D_h \alpha_h^i, \quad (8)$$

where  $D_h$  is the HR dictionary. Conventional sparse representation-based methods assume that the LR patch and its corresponding HR version share the same sparse coefficients in relation to their own dictionaries; namely,  $\alpha_l^i = \alpha_h^i$ . Therefore,

$$x^i \approx D_h \alpha_l^i. \quad (9)$$

HR dictionary is defined as

$$D_h = \arg \min_{D_h} \sum_i \|x_i - D_h \alpha_l^i\|_2. \quad (10)$$

The sizes of the dictionaries  $D_l$  and  $D_h$  are  $D_l \in \mathfrak{R}^{n \times w}$  and  $D_h \in \mathfrak{R}^{m \times w}$ , respectively, where  $w$  is the number of atoms in the dictionary.  $n$  is the dimension of each atom in LR dictionary while  $m$  is the dimension of each atom in HR dictionary.

It is clear that the sparse representation is a bridge between low-resolution and high-resolution patches. To generate such sparse representation, both LR dictionary  $D_l$  and HR dictionary  $D_h$  play a key role. The dictionaries  $D_l$  and  $D_h$  can be easily generated from a set of samples by the methods such as OMP [13].

Once sparse coefficients for each LR patch are learned, we can use this sparse representation to recover its corresponding HR patch. If we have obtained all the reconstructed HR patches, the HR image is recovered by averaging the overlapping reconstructed patches on their overlaps.

### 3. The Proposed SR Method

The proposed method can be divided into three steps: (a) learning different dictionary based on different morphologies, (b) calculating the projection matrixes, and (c) reconstructing the HR mobile sensor image.

*3.1. Learning the Dictionaries Based on Different Features.* Most existing sparse representation-based SR methods use only derivative features to represent the morphologies of LR image. However, the artifacts would occur when using inappropriate features. An explanation for this phenomenon is that dictionary learning from only one kind of features cannot represent essential morphologies of the images. Since the morphologies can vary significantly across images, different patches prefer different features for representation of their morphology accurately. As such, multifeature treatment can help represent the image in a more efficient manner. We propose a method which can present the image with different dictionaries based on different features.

For LR patch  $\text{Patch}_{y^i}$ ,  $K$  different types of features can be adopted to represent it:

$$y^{i,k} = \text{Features}_L^k(\text{Patch}_{y^i}), \quad k = 1, 2, \dots, K, \quad (11)$$

where  $y^{i,k}$  ( $k = 1, 2, \dots, K$ ) is the  $k$ th kind of features of  $\text{Patch}_{y^i}$ .  $\text{Features}_L^k(\cdot)$  denotes extracting  $k$ th kind of features.

Similarly, for the HR patch  $\text{Patch}_{x^i}$ ,  $x^i$  is the feature of it.

Given HR patch  $\text{Patch}_{x^i}$  and LR patch  $\text{Patch}_{y^i}$ , we can obtain  $K$  kinds of LR and HR patch pairs  $\{y^{i,k}, x^i\}$  ( $k = 1, 2, \dots, K$ ) for training.

Based on the  $K$  kinds of LR and HR training sets  $\{y^{i,k}, x^i\}$  prepared above, the LR and HR dictionaries of these training sets are learned from the following models.

The  $K$ -SVD dictionary training is applied to the set of patches  $\{y^{i,k}\}$ :

$$\begin{aligned} D_l^k &= \arg \min_{D_l^k, \alpha_l^{i,k}} \|y^{i,k} - D_l^k \alpha_l^{i,k}\|_2^2 \\ &\text{s.t.} \quad \|\alpha_l^{i,k}\|_0 \leq L \quad \forall i, \end{aligned} \quad (12)$$

where  $\alpha_l^{i,k}$  are sparse coefficient vectors of  $y^{i,k}$  and  $\|\cdot\|_0$  is the  $l^0$  norm counting the nonzero entries of a vector. Most sparse representation-based SR methods rely on the assumption that the HR and LR images share the same set of sparse coefficients. Therefore, the HR image can be reconstructed by combining the HR dictionary and the sparse coefficients of the corresponding LR image. Thus, the HR patch  $x^i$  can be recovered by approximation as  $x^i \approx D_h^k \alpha_l^{i,k}$ .  $D_h^k$  can be calculated by minimizing the following mean approximation error; that is,

$$D_h^k = \arg \min_{D_h^k} \|x^i - D_h^k \alpha_l^{i,k}\|_2^2. \quad (13)$$

*3.2. Calculating the Projection Matrixes.* Although sparse representation-based methods offer a good performance, the

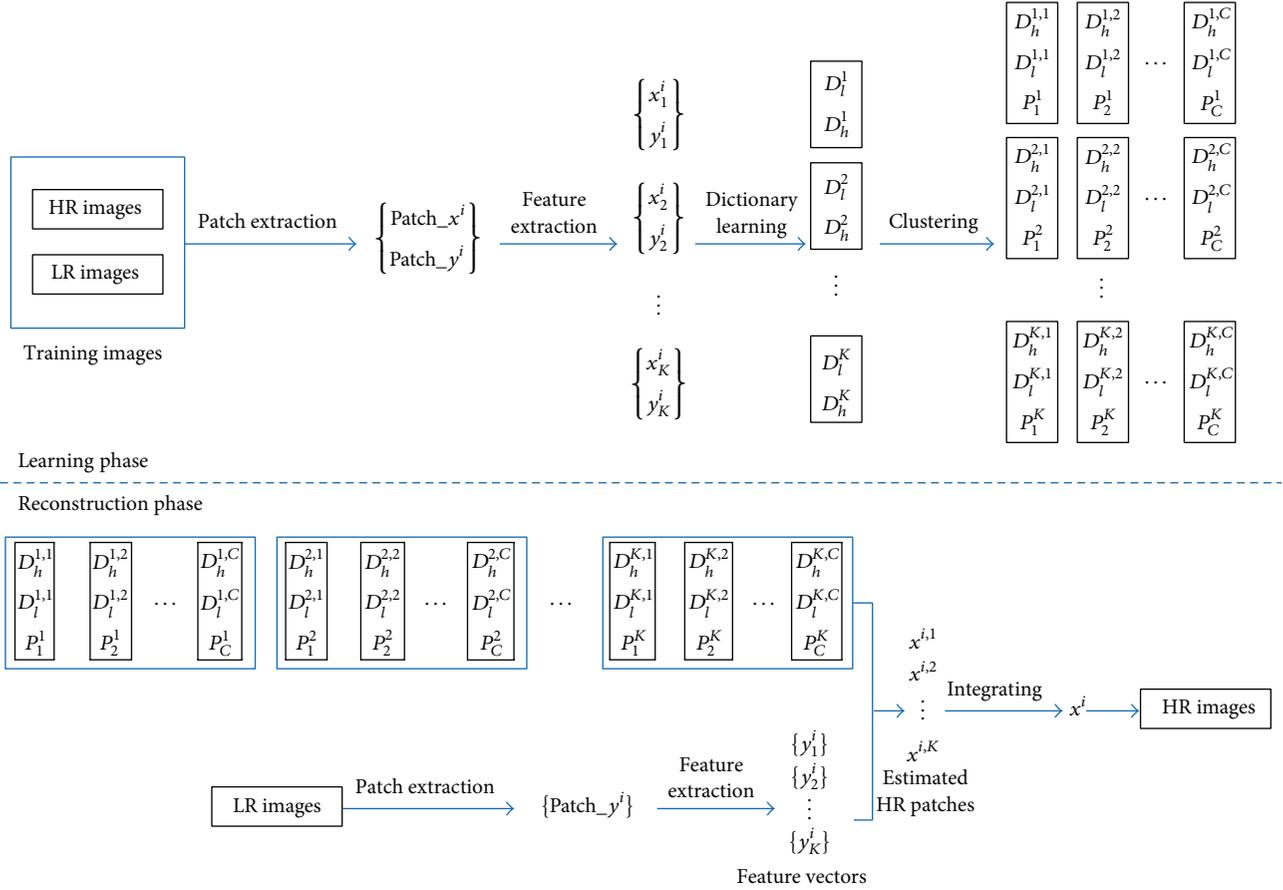


FIGURE 1: Framework of sparse representation-based SR method.



FIGURE 2: Test images in our experiments.

optimization of dictionary learning and image reconstruction has a problem of highly intensive computation. Besides, sparse representation-based SR methods need to reserve memory to store the information of HR dictionary and LR dictionary. The size of dictionary impacts the memory usage. Sparse representation-based SR methods require intensive large memory, especially with increasing size of dictionary. Both time-complexity and memory usage are key limit factors in the embedded system applications of these methods.

Timofte et al. [19] proposed an Anchored Neighborhood Regression method, which constructed a set of mapping matrixes between the LR and HR patches using learned LR and HR dictionaries.

Based on multiple dictionaries obtained in Section 3.2, Anchored Neighborhood Regression method solves this

problem as follows: for each dictionary, to calculate the sparse representation of  $y^i$ , problem (7) is reformulated as a least squares regression regularized by the  $l_2$ -norm of the coefficients [19]:

$$\alpha_l^k = \min_{\alpha^k} \|D_l^k \alpha^k - y^i\|_2 + \lambda \|\alpha^k\|_2, \quad (14)$$

where  $D_l^k$  is the LR dictionary of the  $k$ th type of feature.  $D_h^k$  is the corresponding HR dictionary of  $D_l^k$ .  $\alpha_l^k$  is the sparse vector of  $y^i$ .

Then, Ridge Regression is employed to solve the problem. The algebraic solution [19] is given as

$$\alpha_l^k = \left( (D_l^k)^T D_l^k + \lambda I \right)^{-1} (D_l^k)^T y^i. \quad (15)$$

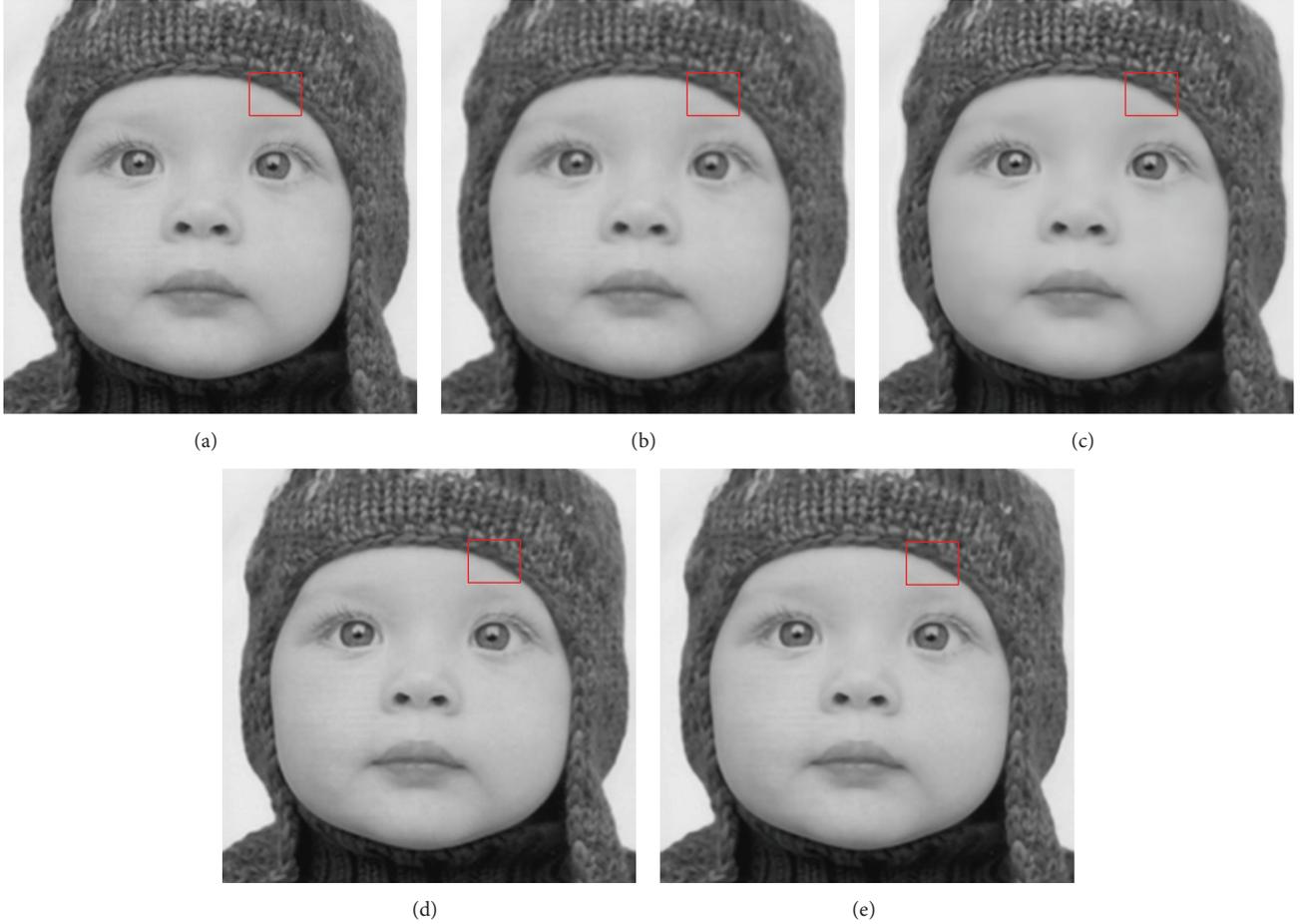


FIGURE 3: Visual comparison of baby: (a) original HR image, (b) results obtained using Cubic B-spline method, (c) results obtained using Yang's method, (d) results obtained using the ANR method, and (e) results obtained using our method.

Since sparse representation-based SR methods assume that the HR and LR images share the same set of sparse coefficients, therefore, the HR patches can be reconstructed by the sparse coefficients of the LR image  $\alpha_l^k$  and the corresponding HR dictionary  $D_h^k$ :

$$x^i = D_h^k \alpha_l^k = D_h^k \left( (D_l^k)^T D_l^k + \lambda I \right)^{-1} (D_l^k)^T y^i. \quad (16)$$

We can obtain mapping matrixes between the LR and HR patches:

$$x^i = P_G^k y^i, \quad (17)$$

$$P_G^k = D_h^k \left( (D_l^k)^T D_l^k + \lambda I \right)^{-1} (D_l^k)^T.$$

Equation (17) means that we can precalculate a mapping matrix for each dictionary. Inferring the HR patch becomes a multiplication for each input patch. The mapping matrix  $P_G^k$  can be computed offline and saved as a simple matrix to be applied to new image patches, which makes vast computational speedup while keeping the same accuracy as previous methods.

Timoft et al. [18] group the dictionary atoms into neighborhoods. More specifically, for each atom in the dictionary, they compute its  $K$  nearest neighbors, which will represent its neighborhood. Once the neighborhoods are defined, Anchored Neighborhood Regression method calculates a separate projection matrix  $P_i^k$  for each dictionary atom  $d_i^k$  based on its own neighborhood. The SR problem can then be solved by calculating the nearest atom in the dictionary for each input patch feature. Then, the HR patch can be reconstructed using the projection matrix of the nearest atom:

$$x^i = P_i^k y^i, \quad (18)$$

$$P_i^k = D_h^{k,i} \left( (D_l^{k,i})^T D_l^{k,i} + \lambda I \right)^{-1} (D_l^{k,i})^T,$$

where  $P_i^k$  is the projection matrix of the atom  $d_i^k$  and  $d_i^k$  is the nearest atom of  $y^i$  in the LR dictionary  $D_l^k$ .  $D_l^{k,i}$  is the neighborhoods set of atoms  $d_i^k$ .  $D_h^{k,i}$  is the corresponding set of the HR dictionary  $D_h^k$ .

Anchored Neighborhood Regression method reformulates the SR problem as a least squares regression, which

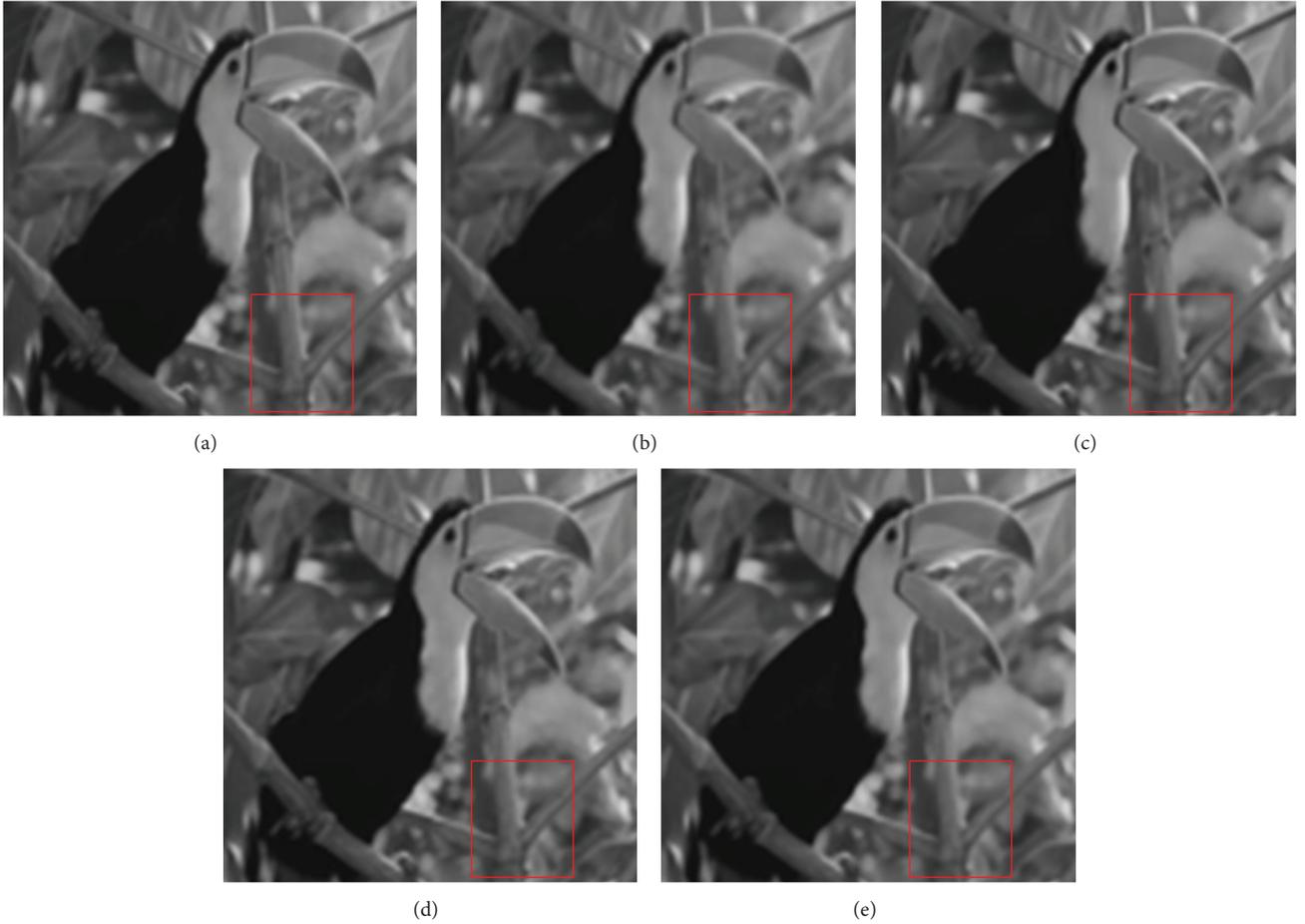


FIGURE 4: Visual comparison of bird: (a) original HR image, (b) results obtained using Cubic B-spline method, (c) results obtained using Yang's method, (d) results obtained using the ANR method, and (e) results obtained using our method.

leads to a vast computational speedup. However, Anchored Neighborhood Regression method still needs to store separate projection matrix for each dictionary atom which is high memory usage. Memory usage is key limit factor in the embedded system application for Anchored Neighborhood Regression methods.

To reduce the memory usage, we propose an Anchored Cluster Regression method. This method divides the atoms into several clusters for each dictionary by  $K$ -means clustering. Then, separate projection matrix  $P_c^k$  of each cluster is calculated. Then, use the projection matrix of the nearest cluster to reconstruct the HR patch:

$$x^{i,k} = P_c^k y^i, \quad (19)$$

$$P_c^k = D_h^{k,c} \left( (D_l^{k,c})^T D_l^{k,c} + \lambda I \right)^{-1} (D_l^{k,c})^T, \quad (20)$$

$$\alpha_l^k = \left( (D_l^{k,c})^T D_l^{k,c} + \lambda I \right)^{-1} (D_l^{k,c})^T y^i, \quad (21)$$

where  $D_l^{k,c}$  is the set of atoms in the cluster  $c^k$  of the LR dictionary  $D_l^k$ .  $D_h^{k,c}$  is the corresponding set of the HR dictionary  $D_h^k$ .

Anchored Cluster Regression method only needs to store the projection matrix of each cluster rather than the projection matrix of each atom. If  $N$  atoms are divided into  $K$  clusters, Anchored Cluster Regression method only needs to store  $K$  projection matrix of each cluster, while Anchored Neighborhood Regression needs to store  $N$  projection matrix of each atom. Anchored Cluster Regression significantly reduces the memory. Furthermore, the computational complexity of Anchored Cluster Regression method is  $O(K)$ , while complexity of Anchored Neighborhood Regression is  $O(N)$ , where  $N$  is the number of atoms and  $K$  is the number of clusters. Anchored Cluster Regression significantly reduces the computation.

**3.3. Reconstructing the HR Image.** Given a LR patch, we can get different HR patches based on different projection matrixes. These different HR patches are integrated to generate the final reconstructed HR image.

For a LR patch  $\text{Patch}_l y^i$ , we get the  $K$  kinds of features  $\{y^{i,k}\}$  ( $k = 1, 2, \dots, K$ ). For the features  $y^{i,k}$ , we can find the nearest cluster  $c^k$  of the  $k$ th LR dictionary; then, we can obtain  $k$ th estimated HR patch  $x^{i,k}$  of the HR patch  $x^i$  based on the projection matrix by (19).

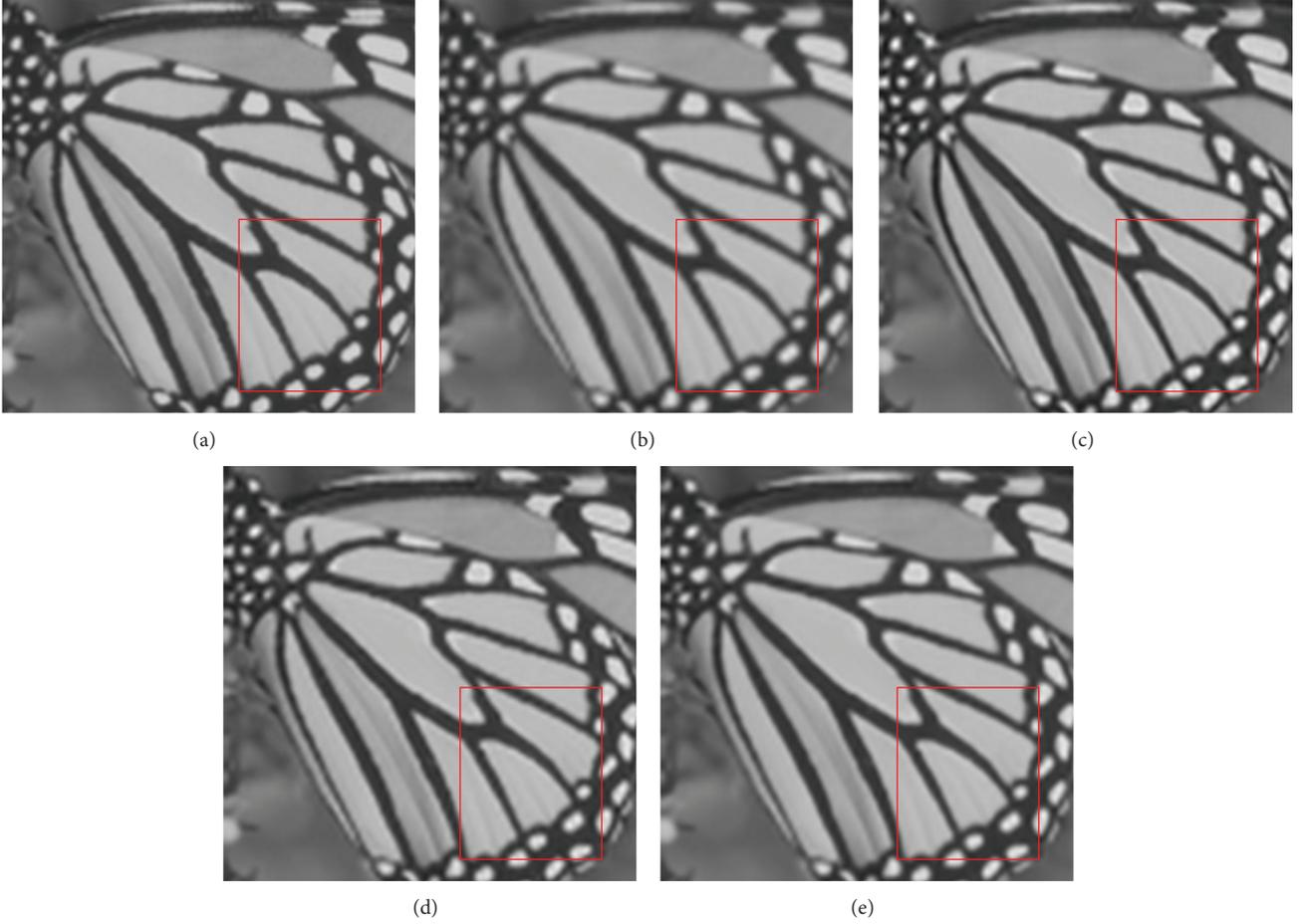


FIGURE 5: Visual comparison of butterfly: (a) original HR image, (b) results obtained using Cubic B-spline method, (c) results obtained using Yang's method, (d) results obtained using the ANR method, and (e) results obtained using our method.

Those  $K$  different estimated HR patches are fused together to get a final reconstructed HR image  $\hat{x}^i$  of HR patch  $x^i$  [20]:

$$\hat{x}^i = \frac{\sum_{k=1}^K w^k x^{i,k}}{\sum_{k=1}^K w^k}, \quad (22)$$

where  $w^k$  is important. According to the weight  $w^k$ , this study fuses the  $K$  different estimated HR patches  $\{x^{i,k}\}$  ( $k = 1, 2, \dots, K$ ) together to get the final reconstructed HR image  $\hat{x}^i$ :

$$w^k = \exp\left(-\frac{(e_l^k)^2}{2\sigma^2}\right), \quad (23)$$

where  $e_l$  is representation error function.  $e_l$  reflects the accurateness of the sparse representation:

$$e_l = y^{i,k} - D_l^{k,c} \alpha_l^{k,c}, \quad (24)$$

where  $e_l$  is smaller,  $D_l^{k,c} \alpha_l^{k,c}$  is more similar to  $y^{i,k}$ .

**3.4. Summary of Proposed Algorithm.** The proposed method contains two phases, that is, learning phase and reconstruction phase. For the learning phase, features of different morphologies are extracted from training images. Then, valid multiple dictionaries are learned based on different morphologies. For each dictionary, the atoms are divided into multiple clusters. The projection matrix of each cluster is calculated by (19)–(21).

In the reconstruction phase, for each LR patch, features of different morphologies are first extracted. Then, for each type of features, the nearest clusters in its corresponding morphology dictionaries are found. Based on the projection matrixes of these clusters, multiple estimated HR patches are reconstructed in the final stage. Then, the final HR patches are generated by using weighting average to process all estimated HR patches. Ultimately, the HR image is composed through averaging the overlapping reconstructed patches. The algorithm is illustrated in Figure 1.

## 4. Experimental Results

In our experiments, we use the same training set as [18], which contains 91 images. For test, we use Set 5 datasets from

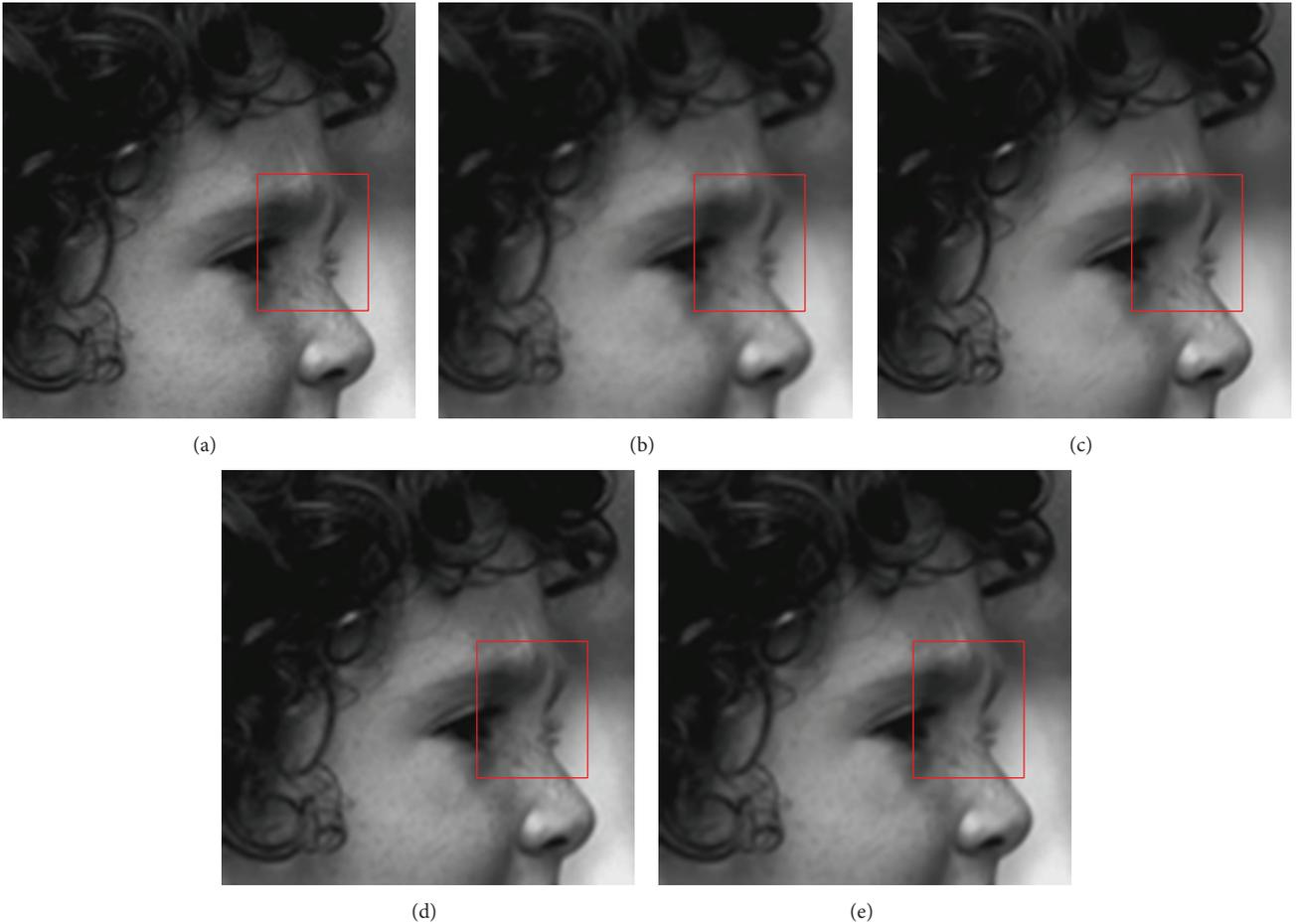


FIGURE 6: Visual comparison of head: (a) original HR image, (b) results obtained using Cubic B-spline method, (c) results obtained using Yang's method, (d) results obtained using the ANR method, and (e) results obtained using our method.

[18]. The Set 5 datasets contain 5 images. The test images are shown in Figure 2. The low-resolution images used in all experiments are downsampled from the high-resolution images. In our experiments, low-resolution images are generated by shrinking the corresponding high-resolution images with the scale factor of 3.

Gabor filters [21] have similar frequency and orientation representations to those of the human visual system. In the spatial domain, a 2D Gabor filter can be generated by Gaussian kernel function modulated with a sinusoidal plane wave. We employed Gabor features to represent the texture structure, since Gabor features can effectively characterize the texture representation and discrimination of LR patch. Gabor features of the image are extracted by convolving the normalized images with a family of Gabor filter (in this study, we use 2 scales and 4 orientations). Beside, we use derivative features representing high-frequency morphology structure of the image.  $k = 1$  means that the features are Gabor features.  $k = 2$  means that the features are derivative features.

In the low-resolution images, we always use  $3 \times 3$  low-resolution patches, with overlap of 1 pixel between adjacent patches, corresponding to  $9 \times 9$  patches with overlap of 3 pixels for the high-resolution patches. 10000 pairs of low- and

high-resolution patches are randomly chosen from the patch pairs generated by training images for dictionaries training.

Not only visual comparison but also quantitative comparisons are confirming the superiority of the proposed method. Peak signal-to-noise ratio (PSNR) and the structural similarity measurement (SSIM) have been implemented in order to obtain some quantitative results for comparison. The values of the PSNR and SSIM of all of the test images were used as the quality index. The PSNR evaluates the reconstruction quality based on the pixel intensity. The SSIM measures the similarity between two images based on their structural information. The SSIM metric needs a "perfect" reference image for comparison and provides a normalized value within  $[0, 1]$ , where "0" indicates that the two images are totally different, whereas "1" confirms that the two images are the same. Thus, higher values of PSNR and SSIM indicate a result with better quality.

*4.1. The Effect of Dictionaries with Multifeatures.* To validate the effect of dictionaries with multiple features, we compared our method with the derivative feature-based method and Gabor feature-based method. In our method, we use Gabor features to characterize the texture of the image and derivative

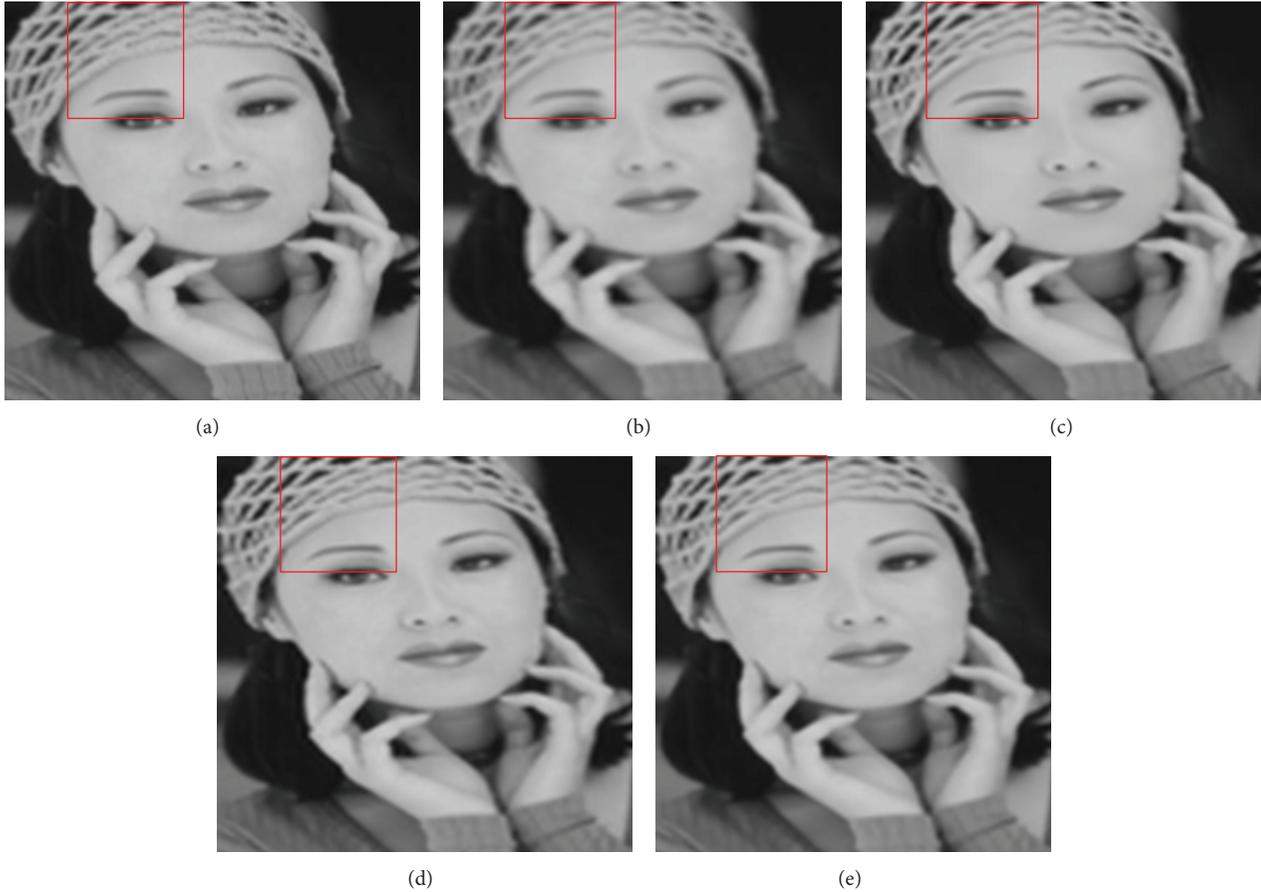


FIGURE 7: Visual comparison of woman: (a) original HR image, (b) results obtained using Cubic B-spline method, (c) results obtained using Yang's method, (d) results obtained using the ANR method, and (e) results obtained using our method.

features represent high-frequency morphology structure of the image. These two features are used together to present the image. The derivative feature-based method adopted derivative features to represent the LR image. And dictionaries with the derivative features were used to sparsely represent the LR patch. Similarly, Gabor feature-based method adopted Gabor features to represent the LR image. And dictionaries with the Gabor features were used to sparsely represent the LR patch. The PSNR and SSIM values of the SR results using various methods are listed in Table 1. We can observe that the proposed framework has a better performance than the single feature-based methods in terms of both PSNR and SSIM. This is due to the reason that one single feature extraction approach cannot accurately capture all the essential features of the image. Multitype features can jointly represent different morphologies in the image more accurately, so that different dictionaries with multitype features can represent the image in a more efficient manner and provide a more global look of the image, which would lead to the fact that the proposed framework has a better performance.

**4.2. Reconstruction Results.** To illustrate the effectiveness of the proposed framework in terms of visual fidelity and objective criterion, we compared the proposed framework with three well-known image SR algorithms, that is, Cubic

TABLE 1: Effects of multifeatures.

Method	Gabor features	Derivative features	Multiple features
	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
Baby	33.67/0.681	34.69/0.731	<b>35.09/0.743</b>
Bird	32.43/0.761	33.16/0.783	<b>34.42/0.802</b>
Butterfly	24.65/0.658	24.58/0.647	<b>25.79/0.680</b>
Head	29.46/0.673	33.19/0.712	<b>33.56/0.728</b>
Woman	24.94/0.578	29.86/0.683	<b>30.16/0.693</b>

B-spline interpolation method, Yang's method [16], and Anchored Neighborhood Regression method [18]. For ANR, in dictionary learning step, the number of atoms in the dictionary is 1,000. The neighborhood size is 40 when regressor is calculated.

Yang's method needs to store the information about the HR and LR dictionary. The sizes of HR patch and LR patch are  $\sqrt{nL} \times \sqrt{nL}$  and  $\sqrt{n} \times \sqrt{n}$ , respectively. If the atom number in the dictionary is  $M$ , with  $t$  iterations needed, the computational complexity of Yang's method is roughly  $O(kMnL)$ . Anchored Neighborhood Regression needs to store  $M$  projection matrix of each atom in the dictionary. The size of projection matrix is  $nL \times n$ . The computational

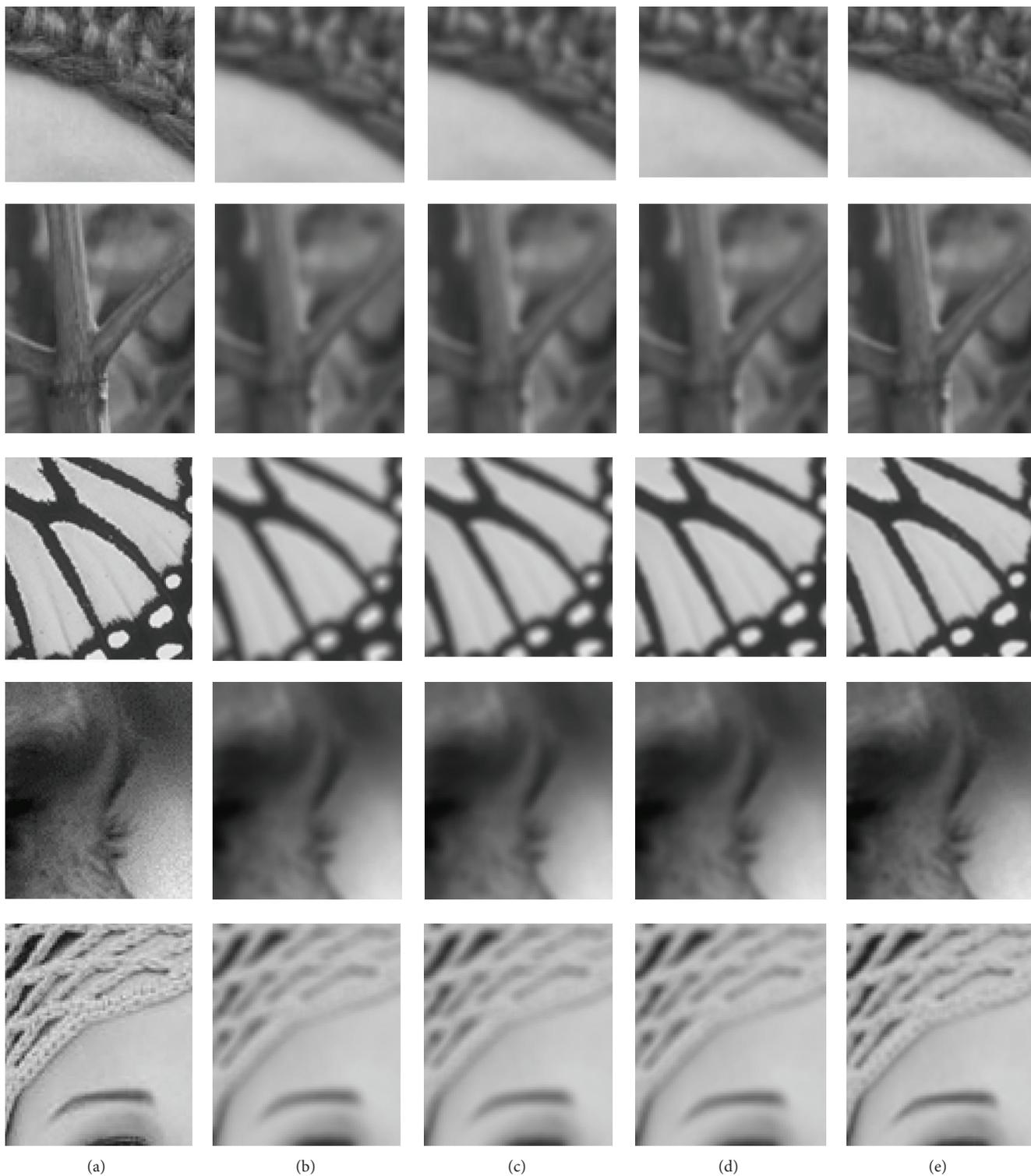


FIGURE 8: Zoomed version of the marked area of Figures 3–7: (a) original HR image, (b) results obtained using Cubic B-spline method, (c) results obtained using Yang's method, (d) results obtained using the ANR method, and (e) results obtained using our method.

complexity of Anchored Neighborhood Regression is  $O(N)$ ; Anchored Cluster Regression method only needs to store the projection matrix of each cluster rather than the projection matrix of each atom. The size of projection matrix is  $nL \times n$ .

If  $N$  atoms are divided into  $K$  clusters, Anchored Cluster Regression method only needs to store  $K$  projection matrix of each cluster. The computational complexity of Anchored Cluster Regression method is  $O(K)$ .

TABLE 2: Numerical results of Figures 3–7.

	Cubic B-spline PSNR/SSIM/time (second)	Yang's PSNR/SSIM/time (second)	ANR PSNR/SSIM/time (second)	Our method PSNR/SSIM/time (second)
Baby	33.91/0.697/0	34.29/0.726/56.06	34.90/0.738/0.98	<b>35.09/0.743/0.8</b>
Bird	32.58/0.768/0	34.11/0.789/22.33	34.38/0.794/0.29	<b>34.42/0.802/0.24</b>
Butterfly	24.04/0.632/0	25.58/0.674/20.22	25.82/0.687/0.24	<b>25.79/0.680/0.20</b>
Head	32.88/0.704/0	33.17/0.709/15.58	33.44/0.716/0.28	<b>33.56/0.728/0.26</b>
Woman	28.56/0.679/0	29.94/0.687/19.04	<b>30.17/0.693/0.27</b>	<b>30.16/0.693/0.23</b>

We first presented the superresolution results by different methods in Figures 3–7. Then, the zoomed version of the marked areas of Figures 3–7 is shown in Figure 8. We see that the Cubic B-spline interpolation methods blurred the edges and lost some dedicated details in the resultant images. Although Yang's method recovered plenty of details, it produced many jaggy and ringing artifacts along with the edges or details. The reason for this is that a single type of features is unable to completely represent various structures of the image. On the contrary, with the proposed framework, images were presented with sharp edges but fewer artifacts. Different dictionaries with multitype features were used to represent the LR image and the weights of different HR patches were also adaptively adjusted in the proposed framework. That is why the proposed framework obtained a better visual quality than all the other three methods.

Moreover, the PSNR and SSIM values of superresolution results using various algorithms are listed in Table 2. We can see the PSNR and SSIM gains of proposed framework over the other methods, which demonstrates that the superresolution results with the proposed framework have better objective quality than other methods in terms of PSNR and SSIM indexes. All this illustrates that the proposed framework is the best one among the compared methods in terms of visual perception and objective quality.

In addition, from Table 2, we see that the time consumption of Cubic B-spline interpolation methods is 0 seconds. That is because Cubic B-spline is an interpolation-based method which is the simplest and the fastest. We can find that Yang's method is consistently the most time-consuming, because it needs to solve a least squares optimization and an iterative convex optimization. Due to reformulating the SR problem as a least squares regression, our method runs a little faster than Anchored Neighborhood Regression method, that is, because it only needs to search the nearest cluster in our method rather than searching the nearest atom in Anchored Neighborhood Regression method. Besides, our method significantly reduces the memory usage, which is suitable for embedded system applications.

## 5. Conclusion

This paper introduces a new SR enlargement method for mobile sensor image. First, to represent mobile sensor images more accurately, complex information in natural images is optimally captured by different morphological components and multimorphology dictionaries learned from corresponding morphological training set. For each dictionary, the atoms

are divided into multiple clusters to calculate the projection matrix. Then, the weights of the HR patches obtained based on different projection matrix are adaptively controlled. Experiments have proved the improvement of the proposed framework in terms of both visual perception and quantitation comparisons with other compared methods. Since the main computation of this scheme is matrix multiplication and the memory usage is low, this method is easy to implement in mobile embedded system applications.

## Competing Interests

The authors declare that they have no competing interests.

## Acknowledgments

The research is sponsored by The National Natural Science Foundation of China (no. 61271330 and no. 6141101009), the Research Fund for the Doctoral Program of Higher Education (no. 20130181120005), the National Science Foundation for Postdoctoral Scientists of China (no. 2014M552357), the Science and Technology Plan of Sichuan Province (no. 2014GZ0005), and the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry.

## References

- [1] R. Kimmel, "Demosaicing: image reconstruction from color CCD samples," *IEEE Transactions on Image Processing*, vol. 8, no. 9, pp. 1221–1228, 1999.
- [2] X. Li, "Demosaicing by successive approximation," *IEEE Transactions on Image Processing*, vol. 14, no. 3, pp. 370–379, 2005.
- [3] H. Shen, L. Zhang, B. Huang, and P. Li, "A MAP approach to joint motion estimation, segmentation, and super resolution," *IEEE Transactions on Image Processing*, vol. 16, no. 2, pp. 479–490, 2007.
- [4] L. Zhang, H. Zhang, H. Shen, and P. Li, "A super-resolution reconstruction algorithm for surveillance images," *Signal Processing*, vol. 90, no. 3, pp. 848–859, 2010.
- [5] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super resolution," *IEEE Transactions on Image Processing*, vol. 13, no. 10, pp. 1327–1344, 2004.
- [6] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super resolution," *IEEE Computer Graphics and Applications*, vol. 22, no. 2, pp. 56–65, 2002.
- [7] H. Chang, D.-Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in *Proceedings of the IEEE*

- Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '04)*, vol. 1, pp. 275–282, July 2004.
- [8] K. Zhang, X. Gao, X. Li, and D. Tao, “Partially supervised neighbor embedding for example-based image super-resolution,” *IEEE Journal on Selected Topics in Signal Processing*, vol. 5, no. 2, pp. 230–239, 2011.
  - [9] X. Wang and X. Tang, “Hallucinating face by eigen transformation,” *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, vol. 35, no. 3, pp. 425–434, 2005.
  - [10] W. Wu, Z. Liu, and X. He, “Learning-based super resolution using kernel partial least squares,” *Image and Vision Computing*, vol. 29, no. 6, pp. 394–406, 2011.
  - [11] M. Aharon, M. Elad, and A. Bruckstein, “K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation,” *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.
  - [12] G. Monaci and P. Vandergheynst, “Learning structured dictionaries for image representation,” in *Proceedings of the International Conference on Image Processing (ICIP '04)*, pp. 2351–2354, IEEE, October 2004.
  - [13] R. Zeyde, M. Elad, and M. Protter, “On single image scale-up using sparse-representations,” in *Curves and Surfaces*, J.-D. Boissonnat, P. Chenin, A. Cohen et al., Eds., vol. 6920 of *Lecture Notes in Computer Science*, pp. 711–730, Springer, New York, NY, USA, 2012.
  - [14] K. Zhang, X. Gao, D. Tao, and X. Li, “Single image super-resolution with non-local means and steering kernel regression,” *IEEE Transactions on Image Processing*, vol. 21, no. 11, pp. 4544–4556, 2012.
  - [15] X. Yang, K. Liu, Z. Gan, and B. Yan, “Multiscale and multitopic sparse representation for multisensor infrared image super-resolution,” *Journal of Sensors*, vol. 2016, Article ID 7036349, 14 pages, 2016.
  - [16] J. Yang, J. Wright, T. S. Huang, and Y. Ma, “Image super-resolution via sparse representation,” *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2861–2873, 2010.
  - [17] Z. Zhan, J. Cai, D. Guo, Y. Liu, Z. Chen, and X. Qu, “Fast multi-class dictionaries learning with geometrical directions in MRI reconstruction,” *IEEE Transactions on Biomedical Engineering*, 2015.
  - [18] R. Timofte, V. De, and L. V. Gool, “Anchored neighborhood regression for fast example-based super-resolution,” in *Proceedings of the 14th IEEE International Conference on Computer Vision (ICCV '13)*, pp. 1920–1927, Sydney, Australia, December 2013.
  - [19] W. Wu, X. Yang, Y. Pang, J. Peng, and G. Jeon, “A multifocus image fusion method by using hidden Markov model,” *Optics Communications*, vol. 287, pp. 63–72, 2013.
  - [20] M. Perd'och, O. Chum, and J. Matas, “Efficient representation of local geometry for large scale object retrieval,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '09)*, pp. 9–16, Miami, Fla, USA, June 2009.
  - [21] A. K. Jain, N. K. Ratha, and S. Lakshmanan, “Object detection using gabor filters,” *Pattern Recognition*, vol. 30, no. 2, pp. 295–309, 1997.

## Research Article

# CRAM: A Conditioned Reflex Action Inspired Adaptive Model for Context Addition in Wireless Sensor Networks

Majid Hussain,<sup>1</sup> Muhammad Farrukh Shafeeq,<sup>1</sup> Sana Jabbar,<sup>2</sup>  
Ali Hammad Akbar,<sup>1,2</sup> and Shahzad Khalid<sup>3</sup>

<sup>1</sup>Department of Computer Science and Engineering, University of Engineering & Technology (UET), Lahore 54890, Pakistan

<sup>2</sup>Alkharizmi Institute of Computer Science, University of Engineering & Technology (UET), Lahore 54890, Pakistan

<sup>3</sup>Bahria University Wireless Research Center, Bahria University, Islamabad 44000, Pakistan

Correspondence should be addressed to Majid Hussain; majidhussain1976@gmail.com

Received 8 October 2015; Accepted 15 February 2016

Academic Editor: Yurong Qian

Copyright © 2016 Majid Hussain et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In Wireless Sensor Networks (WSNs), Context Awareness is typically realized through Context Aware Systems (CASs). Although almost each CAS follows sense-decide-act cycle, the notion of context is hardwired into the applications; that is, when an event is triggered, the sense-decide-actuate cycle runs and performs required actuation. In situations, for instance, whenever the same event is triggered, the cycle produces the same actuation through mechanical use of the same resources, posing the same processing and time. In this paper, we propose CRAM, a context added system in which actuations once performed by the system help it to internally evolve by serving as new contexts. As the system is exposed to more situations overtime, its context repository is enriched through such retrospective contexts, gradually letting it perform internal actuation through improved introspective contexts. This internal actuation leads the system towards the evolution of intelligent processing by reducing the independent function of decision in sense-decide-actuate cycle and merging it with new context. Finally, the system reaches a juncture where recurrence of each event proves to be a stimulus for the system to respond impulsively, through priming memory of introspective contexts, to achieve an imitation of learned reflex action resulting into reduced time and energy expenditure.

## 1. Introduction

Wireless Sensor Networks (WSNs) comprising a large number of interconnected sensing nodes have been the subject of intensive research for the last one and half decades. The growing applications of WSNs in more or less each aspect of human life have made them the up-to-the-minute topical research area. The broad canvas of WSNs novel applications ranges from volcano, glaciers, and aeroground monitoring, reptile tracking, maritime navigations, and studies of human anatomy at much larger scale to the surveillance and defense which conclusively illustrates the omnipresence of WSNs in human atmosphere. Visual Sensor Networks (VSNs)—special featured WSNs are the networks of visual devices, mostly cameras, equipped with enough onboard processing power to support collaborative image analysis. Being highly diverse in operations, VSNs are integrated with such capabilities where local processing controls video and image data acquisition,

removes cross-layer correlations, and aggregates such data to transmit only what is essential. To strike at such optimal trade-off amongst performance and QoS guarantees amidst time synchronization, storage capability, and multicamera collaboration necessitates autonomous reasoning and decision making in unison in a highly distributed manner.

To achieve such autonomous reasoning, VSNs must be empowered with visual context awareness. Regardless of the sensing type, context awareness is the ability of a system, artifact, or service to be aware of its physical environment and to respond intelligently. In particular case of VSNs, visual context aware systems augment sensing devices with such capability that comprehends the user's real time visual perception and defines corresponding interactions with its environs in a particular situation [1]. In order to accomplish the targeted task, these context aware systems follow a sense-decide-actuate and adapt cycle in which they acquire the context, dig out the situation, reason and decide the suitable

actuation, and then adapt to the resulting situation. A context aware system that has gone through multiple actuations in response to multiple and even so repeating events does not learn the pattern of the occurrences of events. It means that when an event is triggered, it is treated afresh and the system always responds by executing the sense-decide-actuate cycle and carries out the requisite task. It is interesting to observe, however, that for scenarios in which the event is reiterated the sense-decide-actuate cycle is executed in its entirety once again. In other words, a context aware system deals with the same events again and again but it does not have the capability to establish an association between such recurrences of events in order to perform adaptation of its internal functionality. This limitation of context aware systems is exacerbated in VSNs because the recurrence of a visual event in itself is always treated as the arrival of a new scene or a video. Since a context has not been predefined for such recurrence of visual events as a metaevent, a context aware system becomes static and performance limited. The same amount of image processing is replicated at each cooperating node during each new occurrence of the same event and the accumulated processing results are considerably higher. This overprocessing resulting in corresponding time and energy expenditures can be restricted through the knowledge, that is, stored context with associated actuation learned by the system based on its prior encounter with the same event. This associative learning as a metaevent context helps the system to respond impulsively against the recurrence of an event emulating a conditioned reflex arc (reflex arc is the impulsive response of a biological system that bypasses the brain through conditioned behavior). Simultaneously, the actuation once performed by the system helps the system to internally evolve by acting as a new context. As the system is exposed to more situations where the same context occurs over time, its context repository is enriched through such retrospective contexts, gradually letting it perform internal actuation through improved introspective contexts which are learnt-and-stored actuations. Finally, it reaches a juncture where a new situation demands minimal external actuation, hence transforming it into a *context added system*, an ultimate imitation of an autonomic system.

In this paper, we propose CRAM, a novel cut-through processing paradigm for contemporary visual context aware systems. CRAM proposes memory-based visual context addition through a learnt reflex arc implementation. The reflex action is realized through a tenon-mortise layered architecture that uses *priming* memory to associate prior experiences with newer contexts. Using an analytical model and an underlying testbed based on VISTA by Jabbar et al. [2], we show that CRAM results in significant reduction in image processing load, provides energy efficiency, and improved compliance to end-user delay bounds and visual accuracy requirements.

The organization of this paper is as follows: Section 2 briefly presents the contemporary related work on the field. Section 3 describes in detail the proposed architecture. Section 4 deals with the memory hierarchy of CRAM and its similitude with human memory. Section 5 shows performance evaluation through simulation based results

while Section 6 presents an analytical model to validate this research work. Section 7 concludes the paper.

## 2. Literature Survey

This section is tetrapartite. The first part presents context and situation. The second part presents context aware systems in general and visual context aware systems in particular. The third part presents building blocks of context aware systems. And finally the last part explores contemporary applications of bioinspired reflex arc on computing systems.

*2.1. Context and Situation.* One of the prime challenges which have been faced by the researchers is to define the context and exhibit it as a different expression from situation. Baldauf et al. [3], Ryan et al. [4], Abowd et al. [5], and Cheverst et al. [6] refer to context as user's identity, current location, environment, and time illustrating an explicitly non-all-inclusive depiction of context. Dey [7] describes the context as the user's focus of attention, emotional state, date and time, location, and orientation as well as bits-and-pieces and people in user's vicinity. It opens up a new omni-dimensional nature of context where sensing, perception, and measurement of sentiments and conjecturing the focus of intention are of elementary concern. Brown [8] defines context as the elements located in the user's environment about which the computer has the information. These sorts of definitions are often too wide-ranging. Conceivably, the most often used definition has been offered by Abowd et al. [9]. These authors refer to the context as "any information that can be used to characterize the situation of entities (i.e., a person, place, or object) that are considered relevant to the interaction between a user and an application including the user and application themselves." It is generally agreed upon that no single definition of the context even in the contemporary research has been reached. Hull et al. [10] attempt to establish relationship between the context and situation with the former being the characteristics of the latter. Since the concept of situation is very closely related to the notion of context, various authors have continued to relate the two of them. Zimmermann et al. [11] suggest that situation can be taken as an instantaneous and structured representation of a part of the context which may be directly compared with the snapshot taken by a camera. Loke [12] remarks that the situation can be viewed as being at a higher level of abstraction than context. Baker et al. [13] describe the situation as "the complete state of universe at an instance of time." Other studies state that sensors can be used to capture the context and construct high level context models of part of the real world. These models can further be used for the recognition of situation and its corresponding reasoning.

*2.2. Context Awareness and Context Aware Systems.* With a background of user's increasing expectations from its environment to adapt to itself, context aware systems have emerged as powerful means to synergize context generating sensors, wireless networks, and the consequent context utilization. From an interactive perspective, as mentioned by Schmidt [14], context aware systems show either of the

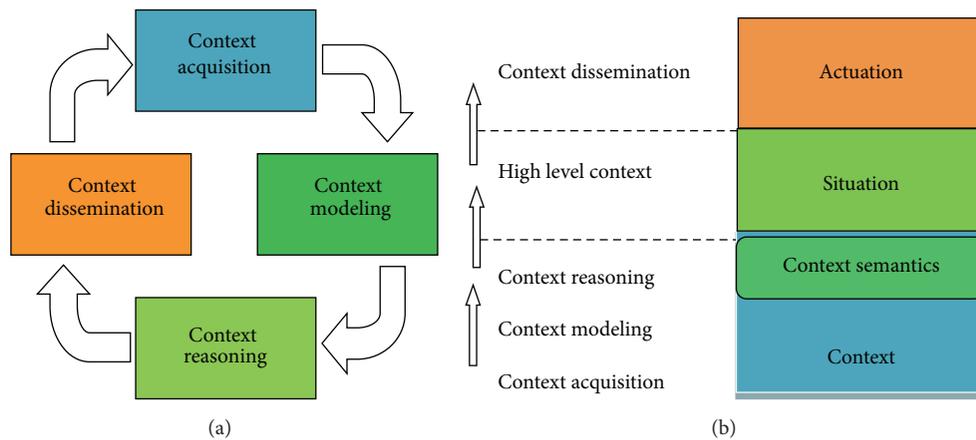


FIGURE 1: (a) Context based building blocks of CAS. (b) Conceptual building blocks of CAS.

two behaviors, adaptive or proactive. Adaptive CASs act on behalf of users, try to adapt to user's context, and are often single trigger based, while, on the other hand, proactive CASs require user's involvement and interaction as they are multitriggered. From an architectural perspective, the CASs can be standalone, centralized, or distributed. Standalone architecture is the simplest and easily deployable but it does not allow internode sharing of context limiting it only for small domain specific applications. In centralized systems, the whole range of sensors and devices is connected to a context server equipped with necessary computing and storage capacity making it simple to add and exchange context offering devices. Chatterjea et al. and Jin and Li [15, 16] propose that while being a single point of failure, the centralized approach is particularly not suited for CASs because the delay involved in context update might change the context itself at the source, whereas Zafeiropoulos et al. [17] take the view that distributed systems offer to use locality-based context acquisition and processing. A comprehensive classification has also been performed by Žontar et al. [18] on the basis of such kinds of architecture.

**2.3. Building Blocks of Context Aware Systems.** A CAS is a complex system that is often defined and composed to meet the desired level of context processing. Usually, context as the pivotal information is the main emphasis of CASs for acquisition, modeling, reasoning, and dissemination. What differs in various approaches is the operational flow and extent of these.

Perera et al. [19] present a cyclic pathway to traverse the context based hierarchy of the building blocks as illustrated in Figure 1(a). Context can be acquired through a diversity of sources. These sources include physical, virtual, and logical sensors. Physical sensors are tangible sources for context acquisition which provide low-level context which is raw, less meaningful, and noisy. Virtual sensors are not direct resources of context but they retrieve diverse types of data from different sources (e.g., phone directories, social websites, profiles, and databases) and publish it in the form of context. The logical sensors (software sensors) combine the physical and virtual sensors to produce high level and more

meaningful context (e.g., weather forecasting web service combines data from temperature, humidity, and wind sensors and corresponding weather maps and seasonal calendars to produce the weather information). Context modeling is the process of defining the context in terms of its characteristics, aspects, quality, and interrelationship with previous contexts and with the set of queries it corresponds to. Context reasoning includes preprocessing, fusion, and inference of context. It handles imperfection and uncertainty of raw data in order to deduce information and in best cases knowledge as high level context [20]. Finally, context dissemination or distribution is its propagation to other entities.

A number of researchers [13, 21, 22] have followed a layered approach to elaborate the fundamental elements of CASs as shown in Figure 1(b). The primary processes of context acquisition, modeling, and reasoning have jointly been incorporated in context and its semantic layers. Context and situation work in cohesion to realize context awareness. Finally, situation as high level context is disseminated for realizing context awareness through actuation layer.

Based upon the above discussed hierarchy, the following CASs are presented. Chen et al. [23] present Context Broker Architecture (CoBrA) to support context aware systems in smart and active spaces. CoBrA maintains a model of current context as *repository of knowledge* shared amongst all the components of the same smart place. The coupling of this shared model with reasoning provides the presence of a user in the smart meeting room. Román et al. [24] come up with a metaoperating GAIA to support the development and execution of portable applications for active spaces. It implements Context File System (CFS) which uses modeling and reasoning by federating application-defined properties and environmental context information to realize active spaces. MobiLife [25] follows similar conceptual framework to realize context aware platform to contact anyone, anytime and anywhere. SPICE [26] demonstrates a tetralayered architecture including a specific layer that they term knowledge layer providing rich set of mechanisms for context acquisition and knowledge processing. Knowledge layer ensures realization of context aware services by making this knowledge available. Open Platform for User-Centric Service

Creation and Execution (OPUCE) [27] emphasizes designing an architecture based on context cycle for integration of communication features in social networking applications and to realize interoperation of Telco-IT applications in a seamless event-driven way. Lamorte et al. [28] demonstrate a similar kind of platform for enabling context awareness in telecommunication services. Baker et al. [13] elaborate the mapping of context life cycle in actual conceptualization of context based situation, perception, decision, actuation, and adaptation process. They further discuss the role of context awareness and CASs based on this hierarchy in future Internet to realize the intelligent society and to address its implications. Barrenechea et al. [29] come forward with a context aware and adaptive approach in distributed event-based systems to model and implement context aware proactive applications involving the combination of context and distributed events. Perera et al. [19] present an IoT paradigm-based detailed analysis of CASs and underscore that each context aware system follows a sense-decide-actuate cycle.

*2.4. Context Aware Systems and Bioinspired Reflex Arc.* It is intuitive to note that there may exist an analogy between context aware systems and reflex actions that are involuntary and automatic responses of living organisms provoked by a sensory stimulus. Although the idea has started to gain strength, there is no contemporary work to the best of our knowledge that truly tries to emulate natural reflex arc in entirety. A recent research project titled “reflex-tree” [30] tries to implement reflex arc for gas pipeline maintenance in urban environments. The authors present a four-tier hierarchy each being part of the reflex arc. The paper however falls short of elaborating how exactly this reflex arc mimics the natural reflex arc. They do not shed any light on the nature of biological reflex arcs if theirs is *instinctive* or *learned*. A deep insight into various studies [31–33] on human and animals provides converging evidence for the existence of two foremost genera of reflexes in human beings: firstly, inborn or intrinsic reflexes which execute their required functions without underlying foundations of memory or prior experience, and these instinctive reflexes are never learned by the subject consciously or unconsciously, and, secondly, *learned* or *conditioned reflexes*, which carry out their requisite operations on the basis of some preceding experience or on the basis of the information items stored in the nondeclarative or implicit unconscious portion of the memory. This specific portion of implicit memory serves as the *priming memory* in which the store is “primed” through repetition of experiences. This primed store helps them to respond very promptly.

On the basis of a thorough review of above-cited literature, our corresponding intuition about context aware systems as sense-decide-actuate cycle, and associative learning-based conditioned reflex action, we conclude the following:

- (1) Each CAS operates through a sense-decide-actuate cycle in which the actuation once performed by the system against a specific event does not help the system to treat the reappearance of the same event intelligently through *on-the-fly* metacontext.

Consequently, CAS executes the same cycle to gain the same actuation resulting in suboptimal performance leading to corresponding time and energy expenditures.

- (2) In contemporary CAS, there is no such mechanism that supports retention and recollection of actuations performed thus far. This limitation restricts the system to adopt the *self*-associative learning ability and therefore cannot react reflexively in case of recurrence of an event.

In this paper, we propose CRAM, a novel cut-through processing paradigm for contemporary visual context aware systems. CRAM provides a cohesive approach to integrate visual context addition and associative learning-based conditioned reflex action to reduce visual context processing so as to evolve CAS into the autonomic state of minimalist actuation. The reflex action is physically realized through supplementation of a memory module into CAS that gets *primed* every time an event occurs. Both context addition and conditioned reflex implemented as an overlay on CAS result in the reduction of node and network-level image processing, increased network-wide energy efficiency, and delay cutbacks.

### 3. CRAM Architecture

In this section, we present conceptualized behavior based upon context addition and reflex arc that forms the basis of the architecture subsequently presented. The layered implementation of the architecture through tenon-mortise is then given.

*3.1. Conceptualized Behavior.* The envisaged behavior of four layers of CASs is shown in Figure 2.

CRAM provides mobile object (MO) detection, tracking, and recognition mechanism through a distributed context added system. The sensing layer being the first and the lowest layer is comprised of pertinent sensing devices. Sensing process takes place whenever a MO is detected. The sensing process always occurs at every node in entirety and it does not undergo any change over time. However, the behavior of other layers changes over time as per Figure 2.

The processing layer that typically implements context semantics in traditional CASs changes its behavior in CRAM with the incorporation of context addition. When MO is detected at the first node, its context is processed in entirety. As MO moves on to the subsequent node in its trajectory, context processing starts to reduce. It is due to the fact that the actuation performed by each predecessor node serves as a metacontext to its successor. The processing and decision-making gradually reach asymptotic minima.

Over time, the corresponding actuation once performed by the system helps the system to internally evolve by serving as a new context. As the system is exposed to more situations in due course of time including recurrence of events, its context repository is enriched through such retrospective contexts, gradually letting it perform internal actuation through improved introspective contexts. Finally, it reaches

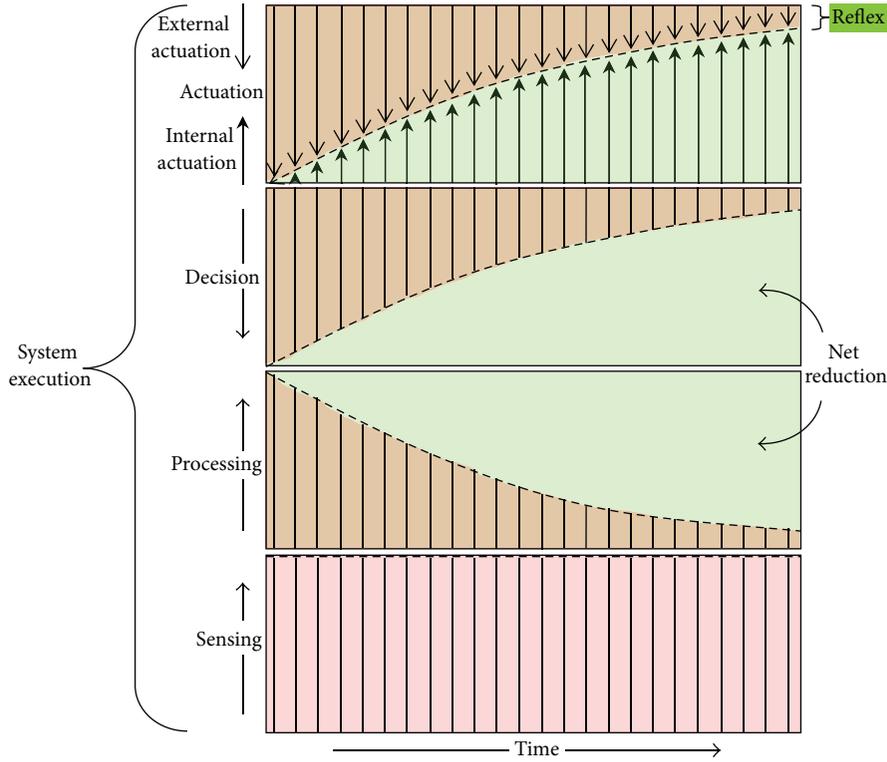


FIGURE 2: Conceptualized behavior of CRAM.

a juncture where a new situation demands minimal external actuation, truly realizing the operation of a *context added system*. When a context added system fully realizes internal actuation, it starts to reflexively respond to recurrence of MO.

**3.2. Tenon-Mortise Architecture.** This section is bipartite. The first part elaborates camera deployment and mobility model that determines the physical topology and resulting mobility pattern of MO in *Region of Interest (RoI)*. The second portion is comprised of layered implementation architecture of CRAM that realizes context addition and reflex arc through cross-layer modules.

**3.2.1. Camera Deployment and Mobility Model.** We formulate the following assumptions for the realization of tenon-mortise layered model:

- (i) The blueprint of RoI is predefined in which each external or edge node (EN) is equipped with sonar and camera while each inner node (IN) is provided with camera only.
- (ii) All SNs have preprogrammed locations in RoI such that each node is aware of its location and the relative locations of each of its one-hop neighbors.
- (iii) All internal nodes have the same computational and memory resources. The external nodes have been provided with an additional memory module that serves as priming (imitation).

TABLE 1: EN states.

State	Sonar	Timer	Camera	Transceiver
Active	1	1	1	1
Quasi sleep	1	0	0	0
Sleep	0	0	0	0

TABLE 2: IN states.

State	Timer	Camera	Transceiver
Active	1	1	1
Demin sleep	0	0	1

- (iv) ENs exhibit three states based upon energy consumption with regard to sonar, camera, timer, and transceiver operations as shown in Table 1, while INs exhibit two states of activation with respect to camera, timer, and transceiver as depicted in Table 2.

- (v) The fields of view of EN sonar and camera are calibrated to be exactly the same.

Since we present a mobile object detection, tracking, and recognition system, it is important to lay out the physical deployment of sonars and cameras and consequent mobility considerations. As MO mobility is constrained to the road segments only (Figure 3), we find Manhattan mobility model as the most suitable one for MO in RoI.

CRAM is comprised of two sensor node stratum. The outer stratum consists of two layers of ENs. The twofold

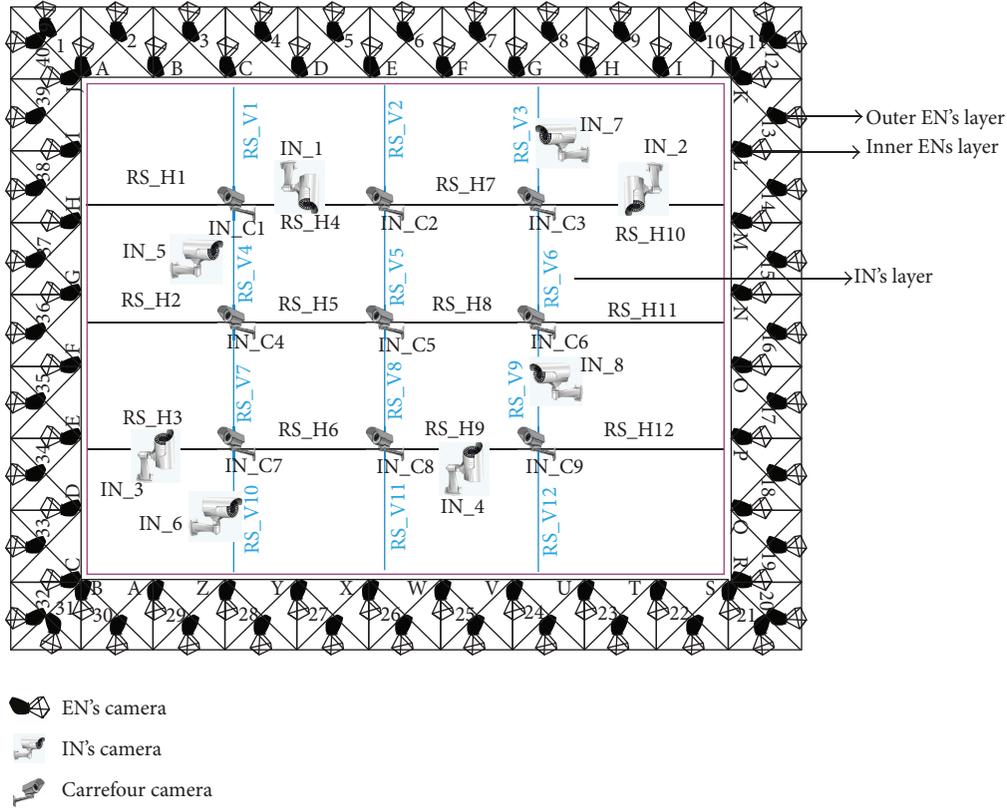


FIGURE 3: Camera deployment and MO mobility pattern.

edge node hierarchy is deployed to assure the detection and recognition of intruding MO in a reliable and energy-efficient manner. The operation of sonars in sets of duo with overlapping (FoV) is presented to ensure coverage while the operation of sonars in triplets achieves energy efficiency.

*Considerations for Camera Deployment of ENs.* For the length of each side of exterior layer as  $L_1$ , the total number of nodes  $N_1$ , each with Field of View (FoV) of width  $W_1$ , required to cover is given by  $N_1 = L_1/W_1$  and the total number of ENs required to secure the complete perimeter =  $4N_1$ .

Similarly, for interior layer,  $N_2 = L_2/W_2$ . The total number of ENs required to secure the complete perimeter =  $4N_2$ .

It is obvious that FoVs of  $W_2 = 2W_1 \Rightarrow N_2 = N_1 - 1 \Rightarrow N_1 = N_2 + 1$ .

*Considerations for Camera Deployments of INs.* In order to detect and track MOs that follow Manhattan model, two types of INs are deployed. The first type of nodes are cameras nodes which are deployed in such a way that each camera covers “ $n$ ” number of horizontal road segments and “ $m$ ” number of vertical road segments in RoI. For instance, as in Figure 3, visual coverage of IN\_1 is three horizontal road segments RS\_H4, RS\_H5, and RS\_H6. Similarly, IN\_8 provides visual coverage of RS\_V7, RS\_V8, and RS\_V9. These road segments all together form rows and columns across RoI as a grid-like structure. The total number of INs required to provide complete RoI coverage is given as  $(IN)_T$ :

$$(IN)_T = (R - 1) + (C - 1), \quad (1)$$

where  $R$  is the total number of rows and  $C$  is the total number of columns.

The second type of INs is Carrefour nodes that are deployed at the Carrefour (Carrefour is French word for intersection). Carrefour nodes provide coverage for MOs approaching or departing an intersection of four road segments. The total number of Carrefour cameras required to cover whole RoI can be given as

$$(IN)_c = (R - 2)(C - 2). \quad (2)$$

*3.2.2. The Layered Architecture.* We present tenon-mortise modules as the realization of the implementation architecture of CRAM based upon cross-layered approach as shown in Figure 4.

Physical layer deals with necessary hardware infrastructure and provides signaling information to the network and processing layers using sonar and timer and shuttered-in frames through cameras. The network layer is just above the physical layer. It implements sonar and timer modules to manage sensing, sleeping, and synchronization operations. It also manages the transceiver through routing decision modules. Memory management module at the network layer allows the management of memory either at the same SN or amongst a set of SNs. Image processing layer, being the highest layer, performs image detection, recognition,

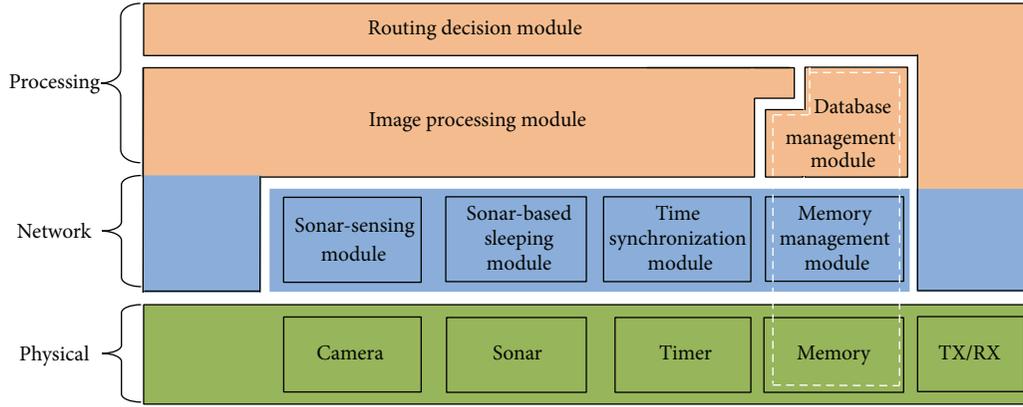


FIGURE 4: Tenon-mortise layered model.

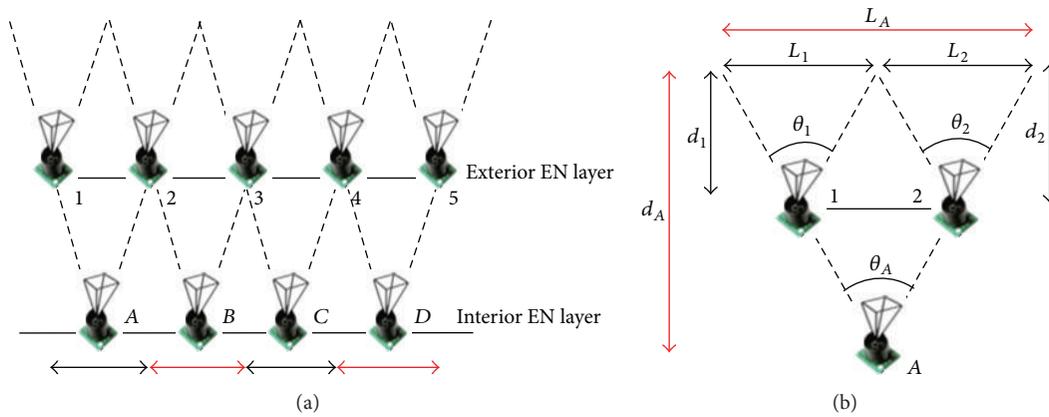


FIGURE 5: (a) Double layer sonar deployment. (b) Triplet formation of sonars.

and tracking through an interplay of database management, image processing, and routing decision operations.

*Physical Layer.* The physical deployment of hardware devices such as sonar, timer, camera, and transceiver is defined at physical layer:

- (i) ENs monitor the existence of an intruding MO through *sonar*. Consequently, cameras are only activated after sonars have detected MO.
- (ii) *Timers* are used to disseminate timing information for synchronization of image acquisition activity at neighboring SNs through in-time activation of next expected nodes.
- (iii) The key function of *camera* is to capture the image of an intruding MO into RoI. The camera is turned on just at the arrival of MO in SNs and is turned off soon after the image acquisition for conserving energy. A multitude of image processing algorithms are then applied.
- (iv) *Memory* provides workspace and storage capabilities. Nonvolatile memory is utilized to store and update topology tables, silhouette tables, and other types of prestored information. Volatile memory provides

run-time environment to perform more intelligence-intensive operations at the upper layers.

- (v) *Transceivers* are used for the transmission and reception of packetized information messages from upper layers.

*Network Layer.* The network layer is comprised of sonar-sensing, sonar-based sleeping, time synchronization, and memory management modules. Memory management module manages the CRAM memory hierarchy with a prime focus to deal with the explicit and implicit memory items. This module ensures the placement of image extracted during the run-time processing of MO in the priming memory that further plays the main role in realization of reflex action (Section 4).

*Sonar-sleeping module* is incorporated with focus on the fact that CRAM assumes to have sonars only at ENs. It becomes improbable to detect MO again in RoI if sonars fail or laxly do so. In order to provide stronger means of detection, two layers of sonars are deployed as shown in Figure 5(a). ENs in both exterior and interior layers are positioned in such a way that one background node resides behind and between two foreground ENs forming a triplet such that FoV of an inner EN is double that of an outer EN (Figure 5(b)).

TABLE 3: CRAM failure resilience through triplet formation.

Malfunctioning triplet	Triplet nodes	Failing nodes	Alternatively working nodes
1	1, 2, A	1 or 2 or 1 and 2 or A	2, A or 1, A or A or 1, 2
2	2, 3, B	2 or 3 or 2 and 3 or B	3, B or 2, B or B or 2, 3
3	3, 4, C	3 or 4 or 3 and 4 or C	4, C or 3, C or C or 3, 4
4	4, 5, D	4 or 5 or 4 and 5 or D	5, D or 4, D or D or 4, 5
5	5, 6, E	5 or 6 or 5 and 6 or E	6, E or 5, E or E or 5, 6

Such triplet formation results in improvements in fault tolerance and failure resilience. For example, when one triplet fails, two neighboring triplets automatically provide alternate coverage. In order to ensure this coverage, intersonar distances of background ENs follow the following well-known solid angle relationship:

$$\begin{aligned} L_1 &= d_1 \theta_1, \\ L_2 &= d_2 \theta_2, \\ L_A &= d_A \theta_A. \end{aligned} \quad (3)$$

As  $L_A = 2L_1$  or  $L_A = 2L_2$ ,  $d_A = 2d_1$  or  $d_A = 2d_2$  such that when IN is on, its coverage should be equal to the coverage of two ENs.

For power consumption analysis, we suppose that the number of ENs at each layer of outer layer is  $n$  (as for very large networks  $n = n + 1$ ). If  $p$  is the amount of power consumed by each node, the average power consumed by outer stratum exterior layer is  $P_{\text{Ext}} = np$ .

Since in each triplet one background node operates against two foreground nodes, the total number of ENs at interior layer for corresponding  $n$  nodes of exterior layer in triplets must be  $n/2$ . Now if  $p$  is the amount of power consumed by each node, then the total amount of power consumed by all nodes of triplets at interior layer can be given as  $P_{\text{Int}} = (n/2)p$ .

If the distance between inner ENs is doubled (requisite for a triplet), the power consumption becomes fourfold. Then, the above relation can be modified as  $P_{\text{Int}} = (n/2)4p \Rightarrow P_{\text{Int}} = 2np$ . And, finally, if the average power consumed by two layers in triplet form is  $P_{\text{Avg}}$ , then it can be demonstrated by the following relation:  $P_{\text{Avg}} = (P_{\text{Ext}} + P_{\text{Int}})/2 \Rightarrow P_{\text{Avg}} = (np + 2np)/2$ .

Hence,

$$P_{\text{Avg}} = 1.5np. \quad (4)$$

So the power consumption difference between nontriplet format and triplet format is  $0.5np$ . Clearly, the triplet formation bears a power tax of  $0.5np$ , but this power added levy provides us with almost complete border breach avoidance system [34] in which the system performs equally well even if 50% of nodes fail. Table 3 shows failure resilience of CRAM through triplet formation.

The power expenditure at ENs can be efficiently managed through a distributed sleep scheduling among the foreground and background nodes of the triplet in such a way that if a foreground EN fails, the background node can be activated

instead. Vice versa, two foreground ENs can be activated in case of background EN malfunction. The sleeping schedule among the foreground and background nodes is shared through SMAC protocol [35] in which background nodes play the role of synchronizers and foreground ones play the role of followers in each triplet.

*Sonar-sensing module* identifies the entrance of the MO in the RoI after receiving Mobile Object Detected (MOD) message from sonar. At the deployment time, distance-based fingerprinting of reflected *Received Signal Strength Indicator* (RSSI) is computed and stored as  $\text{RSSI}_{\text{THRESHOLD}}$  at each EN. The EN at the boundary of RoI periodically sends out beacons to sense possible presence of MO. An EN responds to detection only when the measurement of received RSSI is greater than that of  $\text{RSSI}_{\text{THRESHOLD}}$ . Upon reception of an echo to the beacon at exterior layer EN, the duos of foreground ENs, on the basis of the received RSSI on both of the duo nodes, communicate with each other for the selection of an apt EN to kick off the MO detection job. If the signal is received by the background EN of the triplet, it proceeds with the task of MO recognition itself. For assuring that MO is present and is detected in RoI, we propose that three readings must be taken and analyzed according to Table 4 before camera activation takes place.

Unnecessary camera activation is avoided by anticipating MO trajectory in the network which leads to network longevity. For instance, if a mobile object approaches a sonar and then turns back or moves away from its defined trajectory, no camera is activated. This module, after sensing the existence of any MO in the RoI, generates MOD message to activate the next hop IN.

*Time Synchronization Module.* Each IN sends MO image related information to next hop IN with an associated local time of its acquisition which provides time reference for synchronization [35]. Such synchronization requires accuracy and timeliness so that other INs are activated only and definitely when MO is in this proximity. In order to realize tightly coupled synchronization of INs timer clocks, we present *Camera-Activation Delay Avoidance Time Synchronization* (CADETS) scheme in the time synchronization module (Figure 6) which is tailored to the unique sequence of camera activation and image processing. Here, synchronization activity is initiated in the relevant section of RoI upon the detection of MO by the ENs through beacons to successive INs to update the time information. This time information is further used by INs as reference to synchronize their time clocks and those of subsequent INs in their radio ranges well

TABLE 4: Object movement and EN response.

Sonar reading	MO movement	EN response
$r_1 = r_2 = r_3 = 0$	MO not present	No MO detection by SN
$r_1 = r_2 > r_3$	MO present but moving away	MO detected but not recognized
$r_1 > r_2 > r_3$	MO present but moving away	No MO recognition
$r_1 < r_2 > r_3$	MO present, coming close, and now moving away	No MO recognition
$r_1 = r_2 = r_3$	MO present but standing still	MO recognition is performed at that SN
$r_1 < r_2 = r_3$	Came close and stopped	MO recognition is performed at that SN
$r_1 > r_2 < r_3$	Came close, went away, and came close	MO recognition is performed at that SN
$r_1 < r_2 < r_3$	Moving towards and entering RoI	MO recognition performed all along its trajectory
$r_1 = r_2 < r_3$	Moving towards RoI	MO recognition performed all along its trajectory

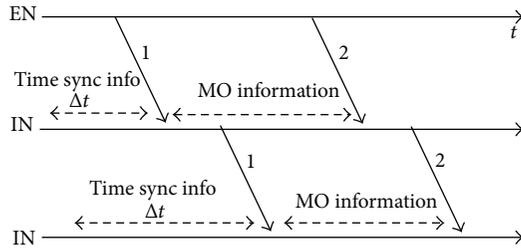


FIGURE 6: Time synchronization through CADETS scheme.

before image processing module disseminates upper layer messages.

*Memory management module* particularly manages the intercommunication of memory module and database management module. One of the main features of this module is to administer the memory module that consists of declarative memory in the form of  $RSSI_{THRESHOLD}$ , topology table, silhouette table, and the compressed images that resulted through SICS. The memory module is also included with nondeclarative memory (priming memory) in the form of priming table comprising aspect ratios of MOs in totality. Memory management module detailed working has been discussed in Section 4.

*Processing Layer.* It contains and executes the logical and algorithmic segment of CRAM to deal with the task of object tracking and identification. It consists of database management module, image processing module, and routing decision module. The operational contribution of each of these modules is elaborated below.

*Database Management Module.* A distributed database based on the regional aspects, application constraints, and requirements to correctly detect, track, and recognize the MO has been deployed over entire RoI.

*Database Organization.* Consider the following:

- (i) Each EN maintains a database record in its memory called topology table to store the neighbor SN's positions and their camera orientations. This table provides the key support in perspective-based mobile object tracking as shown in Table 5.

TABLE 5: SN positions and camera orientations.

SN_ID	SN position (cell, column/row)	SN camera orientation (direction)
IN_1	(1I, I)	W
IN_2	(1, II)	E
IN_3	(12, III)	E

- (ii) Each SN maintains another table in its declarative memory called silhouette table. This table contains silhouettes of probable MOs with their respective identifiers (IDs), silhouettes aspect ratios with respect to their segments, classes assigned to these silhouettes, octets, views, angles, and sureties defined against these stored silhouettes as shown in Table 6. In CRAM, two classes have been assigned to the stored silhouettes. Class 1 represents vehicles while class 2 corresponds to humans. Each silhouette is segmented on the basis of its evident number of prominent visual features and then aspect ratios are computed for each segment of silhouette. This calculation has been discussed in detail at the end of this section. Octet is an ASCII code assigned to a recognized MO. Angle is the angle of camera with respect to MO while the view is the angle of MO with respect to camera. Surety represents the percentage match of acquired image with the stored silhouette and for each stored silhouette it can be given by the following relation:

Silhouette surety at observing SN

$$= \begin{cases} 0 & \text{if no silhouette is present} \\ \frac{1}{k} \times 100 & \text{otherwise,} \end{cases} \quad (5)$$

where  $k$  is the total number of silhouettes of a single object with different aspects distributed throughout the RoI.

- (iii) The static background image of the stored silhouettes is also stored in the database for the further use in the image processing procedures.

*Database Deployment.* Consider the following:

- (i) Topology table deployment: For successful tracking of MO, each SN is equipped with a topology table which

TABLE 6: Silhouette table.

ID	Class	Silhouette	Silhouette aspect ratios with respect to segments		Octet	View	Angle	Surety
			Total segments	Aspect ratios				
1	0 (vehicles)		3	(Segment ID, aspect ratio)	Car	Front	0	12.5%
2			3	(Segment ID, aspect ratio)	Bus	Back	180	12.5%
3	1 (humans)		4	(Segment ID, aspect ratio)	Human	Front	90	25%

TABLE 7: Topology table for IN\_4.

SN_ID	SN position (cell, column/row)	SN camera orientation (direction)
IN_3	(12, III)	E
IN_5	(15, V)	W
IN_6	(10, VI)	E
IN_7	(18, VII)	S
IN_8	(19, VIII)	N
IN_9	(7, IX)	W

contains all neighboring SN's positions and orientations. Table 7 shows the deployment of topology table at IN\_4.

- (ii) Deployment of silhouette table: Storage of silhouette table at each SN depends primarily on its locality. For instance, IN\_1, IN\_2, and IN\_5 are expected to cover a solo side of the MO, so one silhouette of requisite dimension for each expected MO is enough to be stored in its silhouette table as shown through Table 8 for IN\_2. Conversely, as a Carrefour camera is positioned to acquire image through its multiple (sides, front, back, and tilted views) aspects, its table will be provided with all possible dimensions of probable MO silhouettes as shown in Table 9 for IN\_9.
- (iii) It has been erudite that the number of silhouettes in entire network rises with the increase in RoI if the SNs are mounted at a constant distance from each other while if deployed at variable distance, the total number of silhouettes depends upon the number of SNs deployed in RoI.
- (iv) Silhouette table is affected by few factors in which *camera orientation* becomes at primary level. Each SN is stored with the silhouettes which have the higher matching probability with acquired MO silhouettes as for IN\_2 side views of MO have the highest probability to match. Secondly, with the increase in type of MO to be passed through RoI, silhouette table amplifies in its size. Thirdly, in case Gauss's Markov mobility model is used, it will require storing silhouettes of all possible aspects over the entire network and correspondingly

the surety depends upon the number of silhouettes stored for a single object as demonstrated in Table 10.

*Silhouette Segmentation and Aspect Ratio Calculation.* The silhouettes of multiple MOs with the highest matching probabilities are stored in SNs that are used to identify MO by matching them with run-timely acquired ones. Silhouette segmentation process is based on the total number of prominent features; for example, the front view of a human is segmented with five prominent features as head, neck, shoulders, torso, and lower limbs. Figure 7(a) illustrates the different segmented views of a human. Similarly, Figure 7(b) shows the prominent features based segmented images of hatchback and saloon cars with different views.

The computation of aspect ratios is carried out by taking the width-to-height ratios of segmented parts. For instance, Figure 8 explicates the segmentation of human in five parts based on prominent features and calculation of aspect ratios stored in the database table.

*Image Processing Module.* IP module is initiated directly by sonar interruption at EN or upon reception of MOD message at any SN. Being the central part of the CRAM, IP module plays the key role in image capturing and processing. It captures the MO instantaneous image, processes it through different image processing algorithms to convert it into MO silhouette, matches this extracted silhouette with the stored one, recognizes the MO, and presents its outcomes in the form of percentage surety. It operationally proceeds with the following assumptions:

- (i) The prestored and extracted silhouettes are of identical scales.
- (ii) The background subtraction algorithm is restricted to be applied only in the condition when the distance between road segments and the SNs remains the same.

Image processing module contains some submodules whose role is elaborated as follows.

*Image capturing submodule* is responsible for acquisition of instantaneous image. Upon reception of MOD message, the corresponding SN triggers its camera for  $\Delta t$  time and captures fixed size MO image. It is of the fundamental concern that in presence of variance in MO arrival time  $\Delta t$  affects the total "Shutter ON" time. The image acquisition frequency

TABLE 8: Silhouette table of IN\_2.

ID	Class	Silhouette	Aspect ratios with respect to segments		Octet	View	Angle	Surety
			Segments	Aspect ratios				
1	0		3	(Segment ID, aspect ratio)	Bus	Left side	90	12.5%
2			3	(Segment ID, aspect ratio)	Tank	Left side	90	12.5%
3			4	(Segment ID, aspect ratio)	Car	Right side	270	25%
4			4	(Segment ID, aspect ratio)	Jeep	Right side	270	25%
5	1		4	(Segment ID, aspect ratio)	Human	Left side	90	25%

TABLE 9: Silhouette table of IN\_3.

ID	Class	Silhouette	Silhouette aspect ratios with respect to segments		Octet	View	Angle	Surety
			Total segments	Aspect ratios				
1	0 (vehicles)		3	(Segment ID, aspect ratio)	Car	Front	0	12.5%
2			3	(Segment ID, aspect ratio)	Car	Right side	270	12.5%
3			3	(Segment ID, aspect ratio)	Car	Back	180	12.5%
4			2	(Segment ID, aspect ratio)	Bus	Left side	90	12.5%
5			3	(Segment ID, aspect ratio)	Bus	Back	180	12.5%
6			3	(Segment ID, aspect ratio)	Jeep	Right	270	12.5%
7			3	(Segment ID, aspect ratio)	Jeep	Front	0	12.5%
8			3	(Segment ID, aspect ratio)	Car	Tilt	315	12.5%
9			3	(Segment ID, aspect ratio)	Car	Tilt	225	12.5%
10	1 (humans)		3	(Segment ID, aspect ratio)	Man	Left side	90	25%
11			3	(Segment ID, aspect ratio)	Man	Front	0	25%

must be 25 frames per second in vehicle traffic areas [36]. Out of  $n$  frames captured by SN, every  $i$ th frame is processed for MO recognition and tracking. Figure 9 demonstrates the timeline of SNs sequential camera activation for  $\Delta t$  seconds.

After acquisition by image capturing submodule, the captured image is processed through the *image change detection module* to find any change in prestored image with static background to detect the presence of MO. We use Gaussian Mixture Model (GMM) for this change detection for given number of Gaussian components for “evolving” background. The process of change detection can be influenced by

a number of factors at each node which include swaying background objects, slow moving foregrounds, and shadowing or illuminating of light sources with their localized distinctiveness. This emphasizes the adaptation of the image change detection module on each IN to become more sensitive in case of activation of localized background process and less sensitive otherwise. More sensitivity leads to counter the effects of active backgrounds processes through higher number of Gaussian components in mixtures of Gaussians and through correspondingly lesser number when background is more stationary. However, the utilization of large number

TABLE 10: Silhouette table with different aspect ratios stored in database.

ID	Class	Silhouette	Aspect ratios with respect to segments		Octet	View	Angle	Surety
			Segments	Aspect ratios				
1			4	(Segment ID, aspect ratio)	Car	Front/left	45	Depends on stored silhouettes
2			1	(Segment ID, aspect ratio)	Car	Left side/top	90	
3	0		3	(Segment ID, aspect ratio)	Jeep	Left side/tilt	90	
4			3	(Segment ID, aspect ratio)	Car	Right side/back	225	
5			3	(Segment ID, aspect ratio)	Car	Right side/tilt	270	

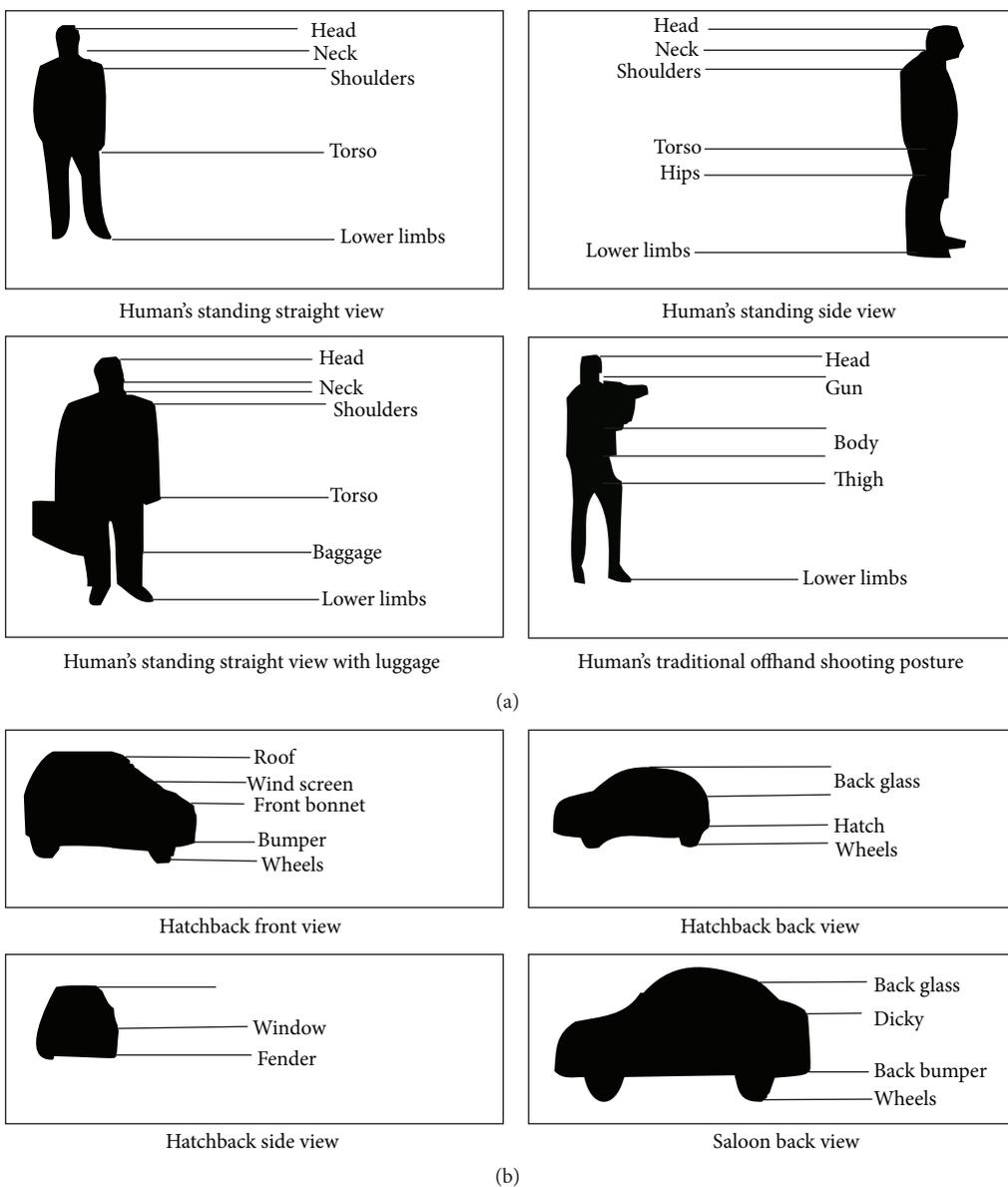


FIGURE 7: (a) Human segmentation with different views based on prominent features. (b) Car segmentation with different views based on prominent features.

TABLE 11: SICS surety based transcoded image storage and corresponding power consumption analysis.

Surety level (%)	Image transcoding level	Image quality level (%)	Power consumption in transcoding operation (mW)
0–25	Level 0	100	Nil
26–50	Level 1	75	68.8
51–75	Level 2	50	44.1
76–99	Level 3	25	24.7

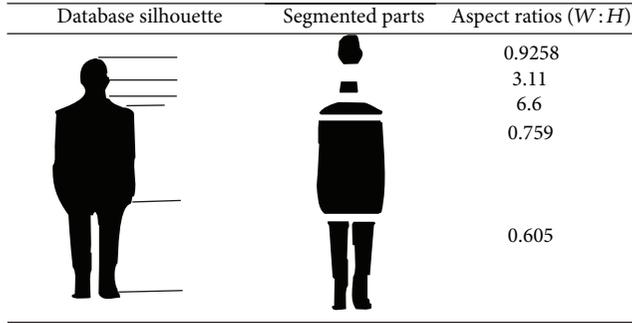


FIGURE 8: Aspect ratios calculation of human silhouette.

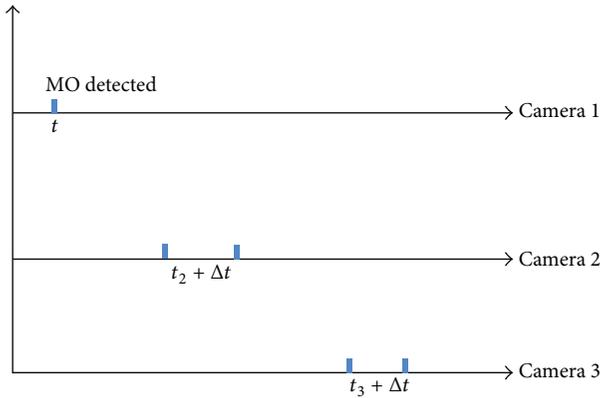


FIGURE 9: Timeline of camera activation.

of Gaussian components is not suggested in CRAM because of higher quantity of consequential energy drain with added complexity.

For optimization of image change detection, we propose an adaptive change detection scheme which is driven by a feedback loop bridged between an up-trajectory and down-trajectory node. The up-trajectory node refers to an EN which activates an IN or an IN which activates another IN while down-trajectory node stands for an IN which is activated by an IN or EN. In the operational execution of this scheme, when an object is sensed by an up-trajectory node, it forwards a MOD message to its down-trajectory neighbor expecting that its counterpart shall detect it. On successful detection, the down-trajectory node replies with positive feedback implying that the change detection at up-trajectory node is sufficiently sensitive or adequate number of Gaussian components are being used at it. In a situation where the down-trajectory node does not detect the MO, it responds

negatively signifying the fact that up-trajectory node is not aptly sensitive or a higher number of Gaussian components are needed to be fuelled. These false positives can also be triggered by malfunctioning of an IN or Carrefour node and can be alleviated by practicing a consensus and voting algorithms among a group of neighboring inner nodes-based hysteresis loop which intertwines recursively in whole network to avoid such malfunctions [37].

The image captured by image capturing module is refined and fine-tuned by image change detection module and is compressed and stored by *Image Compression and Storage Module*. In order to optimize the compression and storage, we use Quality-Aware Transcoding [38] which offers a quality-versus-size trade-off based scheme for dynamically changing the image size. We further propose a compression and storage scheme in which the image is compressed and stored based on its corresponding MO's surety level named as *Surety Based Image Compression and Storage (SICS)*. In this scheme, as a MO attains more surety levels, its image is transcoded to more elevated levels and is stored at lower image quality levels to diminish the energy expenditure as the power consumption of transcoding operation decreases with decline in image quality level. We examine that as the low-quality stored images are sufficient enough for further image processing and MO recognition, the MO recognition process is not manipulated by this low-quality image storage. Moreover, these low-quality stored images occupy less space in the memory which ultimately consumes less computational power for further image processing and MO recognition processes. We justify the proposed method through Table 11.

When a MO penetrates into the RoI for the first time, its image is not transcoded as per the suggestions of SICS and is stored with its original size for better identification. Subsequently, when it passes through more SNs hops (IP) in RoI, it consequently acquires elevated surety levels. At a situation when the MO attains surety level more than 25%, it is assigned with the corresponding transcoding level. For instance, at 50% surety, it achieves transcoding level 1, at 75% surely, it is transcoded at level 2, and when it reaches more than 75% surety, it gets transcoded level 3 and is stored at 75, 50, and 25 percent image quality levels, respectively. The entire power expenditure through image capturing, compression, and storage modules can be given as

$$P_T = P_A + P_C + P_S, \quad (6)$$

where  $P_T$  is the total power consumed while  $P_A$ ,  $P_C$ , and  $P_S$  are powers consumed by image capturing, compression, and storage modules, respectively. Each of these constituents

of power consumption is independently determined by the algorithms used below.

*Image subtraction submodule* is used to extract the MO silhouette by subtracting the silhouette from its image with static background stored at SN through background subtraction submodule. Background subtraction can be optimized in terms of power and time cost by applying Don't Care operation on background image without affecting silhouette extraction. In order to apply Don't Care operation on selected parts of background image, silhouette information of acquired-and-then-stored image is required. The changed and unchanged regions can be then detected by applying the following equation as suggested by Xu et al. [39]:

$$Dx_{ij}^k = x_{ij}^k(t_2) - x_{ij}^k(t_1). \quad (7)$$

The pixel-by-pixel change detection process is executed once at first time entry of the MO in RoI while the change is detected on successive hops. As RoI is based on Manhattan model, the possible division of background image in four portions is proposed in which few are Don't Cared and few are used for background subtraction. It is found that there possibly exist four cases which include the variation of change detection regions and positions of changed regions as shown in Table 12. Based on the total number of changed regions and their positions, we assign a distinctive code to each possible combination which is further used to transfer background subtraction information to neighbor SNs. The background subtraction operation is applied to the one, two, or all parts at all SNs in which change is detected. We also suggest ID-based split background image subtraction which applies the background subtraction operation on some portions of background image on the basis of ID of previous SN. These portions are selected on the basis of scene entry region information where scene entry information refers to the mobile object navigation on last SN where the MO is lastly seen and provides the entry direction of MO at a SN. SNs store this scene entry region information in their topology tables with reference of their neighbor SNs IDs. We present the IN\_4 from Figure 3 as an example to demonstrate the adequate section of background image approaching it from neighbor SNs while supposing that the change is detected in one portion so the subtraction shall take place at corresponding single portion (Table 13).

Further analysis shows that the total time consumed during the image subtraction of one portion is 4.5 times less than the time it takes to subtract whole image.

*Silhouette comparator submodule* compares the finally extracted silhouette with prestored silhouettes in the silhouette table. In order to optimize energy utilization, we present feature-dependent silhouette segmentation (FRILL) procedure. In this segmentation technique, the acquired silhouette is segmented in correspondence with the stored silhouette such that both possess absolutely identical and the same number of segments. Further, the aspect ratio of each segment of extracted silhouette is computed and is compared with the aspect ratios of corresponding segments

TABLE 12: Total possible changed regions with their positions.

Case number	Total number of changed regions	Code number	Changed region
1	4	0000	
2.1	1	0001	
2.2	1	0010	
2.3	1	0011	
2.4	1	0100	
3.1	2	0101	
3.2	2	0110	
3.3	2	0111	
3.4	2	1000	
4.1	3	1001	
4.2	3	1010	
4.3	3	1011	
4.4	3	1100	

of stored silhouette. The silhouette comparison procedure is done through the following expression:

$$\text{Diff} = \text{Min} \left( \sum_j \text{AR}_e^i - \text{AR}_s^i \right) \forall_k, \quad (8)$$

where  $\text{AR}_e$  is aspect ratio of the extracted silhouette,  $\text{AR}_s$  is aspect ratio of the stored silhouette,  $j$  is total number of silhouette segments, and  $k$  is total number of silhouettes deployed on a SN.

Intuitively, this relation works out the extracted and stored silhouettes' aspect ratio difference one by one and then returns the least difference value. The stored silhouette with which it shows minimum difference is declared similar to

TABLE 13: Split background table of IN\_4 based on previous SN.ID.

Previous Node_id	Scene entry region information	Background section
IN_3	Northwest	
IN_5	Southeast	
IN_6	Northeast	
IN_7	Southeast	
IN_8	Southwest	
IN_9	Northwest	
IN_10	Northeast	
IN_11	Northwest	

extracted one while the extent of similitude in percentage is calculated through the following relationship:

$$\text{percentage match} = \frac{\sum_{i=0}^{\text{total segments}} (\text{AR of extracted silhouette})}{\sum_{i=0}^{\text{total segments}} (\text{AR of stored silhouette})} \left(\frac{1}{k}\right) \times 100. \quad (9)$$

The proposed mechanism is presented with an exemplar scenario in Figure 10. The silhouette of an unknown MO is extracted and matched up with four stored silhouettes of human with dissimilar views. The run timely harvested silhouette is segmented in the same correspondence with stored ones in the database. For instance, when this is matched up with standing man's side view, it is segmented in six parts and consequently in five parts while being compared with man's armed view with the same width and height ratios.

After getting done with the matching up task, a packet is generated by IP module in which all the fields are extracted from database except surety as illustrated in Figure 11.

**Routing Decision Module.** This module destines the packets generated by IP modules. The sonar-sleeping, sonar-sensing, time synchronization, and image processing modules are responsible for invoking it. Table 14 shows a variety of decisions this module can take.

When the routing decision module is summoned by sonar-sleeping module, it disseminates the background EN's

TABLE 14: Decision table of routing and decision module.

External module	Packet type	Destination decision
Sonar-sleeping module	Sleeping schedules	Send to ENs in a triplet
Sonar-sensing module	Sleeping schedule	Send to next hop only
Time synchronization module	Beacon/time stamps	Send to subsequent node only
Image processing module	MO recognition information	Send to next hop SN/base station

generated sleeping schedules with destined foreground ENs in a triplet. It further decides about the sonar-sleeping module generated MOD destination where the MOD calls up the camera of most optimal SN for timely acquisition of image frame and MO recognition.

#### 4. CRAM Memory Management and Processing

In this section, we present a mapping between human memory structure and CRAM memory composition. The humans are in-built with a three-stage memory formation as illustrated in Figure 12. The first stage of memory in human anatomy that interacts with the outside world through the sensory receptors is sensory memory.

In CRAM, analogous to haptic and iconic sensory memory, we use sonar as an electromechanical receptor for sensing of a MO in RoI. As discussed in sonar-sensing module (Section 3.2.2), sonar sends beacons continuously in the requisite region but reacts only when the MO presence is detected on the basis of the  $RSSI_{\text{THRESHOLD}}$  information to either commit the event for further processing or pay no heed otherwise.

Short term memory is the second stage of human memory structure which holds items that are of further interest extracted from sensory memory. CRAM uses similar approach for short storage when SICS-based transcoded images are extracted by Image Compression and Storage Module (Section 3.2.2).

For the permanent storage of items, the human anatomy defines long term memory. Long term memory is further subdivided into declarative (explicit) and nondeclarative (implicit) memories. The declarative memory or explicit memory refers to those memories which can be consciously recalled such as facts and knowledge [40]. Declarative memory is further comprised of semantic and episodic memories [41]. Episodic memory is a major constituent of declarative memory that is the collection of previously experienced events with their incidence at particular place, time, coupled emotions, and additional context to figuratively remember the event that took place at certain time and place [42]. Analogous to episodic memory in humans, CRAM uses prestored topology and silhouette tables with their IDs, positions, camera orientations, classes, octets, views, angles,

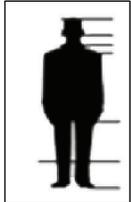
	Database silhouettes	Extracted silhouettes	Segmented parts	H & W of segmented parts	Diff Eq working	
					AR of extracted silhouette	$AR_e - AR_s$
1	 Man's stooping view			Part 1 H = 0.09'' Part 1 W = 0.13''	1.44	1.44 - 2.22 = 0.78
				Part 2 H = 0.03'' Part 2 W = 0.12''	4	4 - 4.16 = 0.16
				Part 3 H = 0.09'' Part 3 W = 0.29''	3.2	3.2 - 3.14 = 0.06
				Part 4 H = 0.29'' Part 4 W = 0.33''	1.1	1.1 - 1.07 = 0.03
2	 Man's standing straight view			Part 5 H = 0.12'' Part 5 W = 0.34''	2.83	2.83 - 1.84 = 0.99
				Part 6 H = 0.32'' Part 6 W = 0.24''	0.75	0.75 - 1 = 0.25
				$(\sum AR_e^i - AR_s^i) = 2.27$		
				Part 1 H = 0.1'' Part 1 W = 0.12''	1.2	1.2 - 1.33 = 0.1
3	 Man's loaded (baggage) view			Part 2 H = 0.05'' Part 2 W = 0.12''	2.4	2.4 - 1.55 = 0.8
				Part 3 H = 0.24'' Part 3 W = 0.03''	8	8 - 8 = 0
				Part 4 H = 0.39'' Part 4 W = 0.3''	0.76	0.76 - 0.815 = 0.05
				Part 5 H = 0.39'' Part 5 W = 0.26''	0.66	0.66 - 0.815 = 0.15
$(\sum AR_e^i - AR_s^i) = 1.1$						
4	 Car's side view			Part 1 H = 0.08'' Part 1 W = 0.14''	1.75	1.75 - 1.33 = 0.42
				Part 2 H = 0.05'' Part 2 W = 0.13''	2.6	2.6 - 1.6 = 0.94
				Part 3 H = 0.05'' Part 3 W = 0.3''	6	6 - 5.16 = 0.84
				Part 4 H = 0.37'' Part 4 W = 0.35''	0.94	0.94 - 1.02 = 0.08
$(\sum AR_e^i - AR_s^i) = 2.55$						
$(\sum AR_e^i - AR_s^i) = 7.19$						
$Diff = \text{Min}_j (\sum AR_e^i - AR_s^i) \forall_k = 1.1$ Result = match with human straight view						

FIGURE 10: Matchup of an unknown MO extracted silhouette with the prestored silhouette in database.

Sequence number	SN_ID	Octet	Angle	Percentage match	Time stamp
-----------------	-------	-------	-------	------------------	------------

FIGURE 11: Resultant MO information packet generated by IP module.

TABLE 15: System specifications for performance evaluation.

Features	Specifications
Matlab (version)	R2012b
Processor	Intel corei5
Memory	4 Gb
Camera	Web camera Logitech Tessar
Processor speed	1.80 GHz
Dimension of the image	240 × 240 pixels
MO traversal (in hops)	12

and sureties. This information is recalled and utilized with timing information as in CADETS through image processing module (Section 3.2.2). The resulting surety and identified class of the extracted image of MO are disseminated for actuation. Concomitant to the operation in the episodic (declarative) part of the memory, aspect ratio of the extracted silhouette becomes the item of interest to be subsequently passed on to and used by nondeclarative memory (implicit).

CRAM exploits the presence of nondeclarative memory in which previous experiences aid the performance of a task *without* conscious awareness of these previous experiences. Such implicit memory is *priming memory* in which the store is “primed” either through repetition of experiences called imitations or the most recent experience and it lets humans respond very promptly [43, 44]. As mentioned in the preceding discussion on episodic memory, CRAM simultaneously primes the nondeclarative memory (priming memory) with the aspect ratio of extracted silhouette. When a new MO is detected, the aspect ratio of its silhouette is compared with the stored aspect ratio of the previous MO in the priming memory. In case of a match, a recognition message is disseminated to the sink without further activating the downstream nodes. Such is the realization of reflex arc that lets the first node observing MO respond while only *referring* to the priming memory as shown in Figure 13.

Since, at the same time, episodic memory is also in the process of computing and comparing aspect ratios of individual segments of the newly detected MO, the results of declarative (episodic) and nondeclarative (priming) memories are evaluated for determining the margin of error. In case of discrepancy, nondeclarative memory processing may be adjusted to yield closer results to those of declarative memory.

## 5. CRAM Performance Evaluation

In this section, we evaluate the performance of CRAM with respect to context addition and reflex arc through prototype implementation and compare its performance with VISTA. It is important to note that an end to end comparison can only exist between VISTA and CRAM because of the common visual context cycle both of them execute. Table 15

TABLE 16: Parameters and situations for experiments.

Parameters	Situations
Lighting effects	Sunlight, artificial light, and shadow
Object types	Human, cars
Angle	Straight, side, and loaded
Background types	Uniform, nonuniform

shows the system specifications and features of the testbed. Table 16 outlines the parameters and situations under which the images were acquired.

**5.1. Reflex Performance of CRAM.** In order to analyze the context added behavior and impulse response of CRAM, a trajectory of 12 hops was implemented for MO. For the first appearance of MO, CRAM demonstrates context addition throughout the trajectory right from the second node, whereas, in case of reappearance of the same MO at the same node, CRAM takes the decision only on the basis of priming memory at the very first node illustrating an obvious imitation of learned reflex action. It is further observed in the particular case of reappearance that CRAM takes 11 times lesser time than VISTA (Table 17). It is due to the fact that VISTA does not declare its recognition results before complete processing of episodic memory for all the nodes in the trajectory.

**5.2. Accuracy of CRAM.** With the implementation of priming memory, it becomes important to be assured on the quality of recognition that reflex arc provides. Table 18 shows that, for multiple iterations of MO recognition, CRAM gave 80% accuracy, the same as VISTA. In both cases, 1 and 4 CRAM erred due to the presence of background objects adding noise. It may also be noted that CRAM performs better than VISTA in poor lighting conditions (case 9) because each preceding process in episodic memory adds its own image processing noise for the succeeding process.

## 6. CRAM Analytical Model

The analysis presented here develops an understanding of context aware systems with regard to the execution of processing and delays incorporated, etc. This section analyzes these aspects with the inclusion of context addition and reflex arc.

**6.1. CRAM Provides Sublinear Bounds on Delay.** In order to verify the time sensitive and delay effective approach of CRAM, we presume a trajectory of  $n$  nodes for a detected MO, where  $N = \{N_1, N_2, \dots, N_n\}$  as shown in Figure 14. Each node  $N_i$  contains a set  $S_i$  whose elements are pieces of silhouette information ranging up to  $m_i$  elements, for example, silhouette, aspect ratio, octets, angles, and surety values.  $A_i$  is subset of  $S_i$  whose elements range from 1 to  $\ell_i$ . Then,

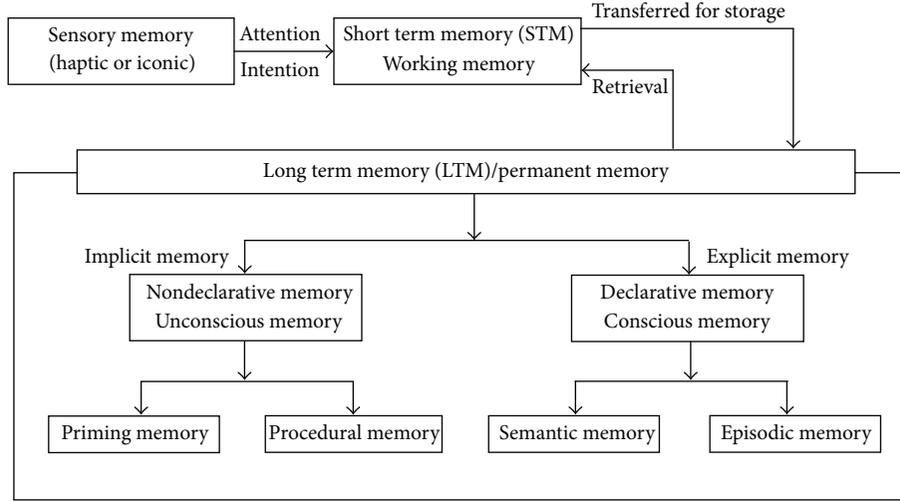


FIGURE 12: Human memory structure.

$$\begin{aligned}
 S_i &= \{1, 2, 3, \dots, m_i\}, \\
 A_i &= \{1, 2, 3, \dots, \ell_i\}, \\
 &\text{where } A_i \subset S_i, \ell_i \leq m_i.
 \end{aligned} \tag{10}$$

*Time Taken at the First Node.* Suppose, at node  $N_1$ , percentage surety level of a detected MO is calculated as a value  $V_1$ . For the calculation of  $V_1$ , the set  $S_1$  of silhouette information at  $N_1$  is traversed for comparison. Consider that a unit time  $t$  is consumed for each comparison.

The best-case time complexity is simply  $t$  if the detected MO matches with silhouette information of  $N_1$  in its very first comparison. It is the realization of reflex arc for recurrent systems where the elements of  $A_i$  are repetitive for successive MO detections.

The average-case time complexity is

$$\frac{1}{m_1} \sum_{i=1}^{m_1} t_i = \frac{m_1 + 1}{2} t, \tag{11}$$

where  $m_1$  is the last object of set  $S_1$ .

The worst-case time complexity of comparison is

$$t \sum_{i=1}^{m_1} 1 = m_1 t \tag{12}$$

if the detected MO matches with silhouette information of  $N_1$  in its very last comparison.

*Time Taken at the Second Node.* At node  $N_2$ , the silhouette information and surety value  $V_1$  calculated by node  $N_1$  is received as a new context according to our context addition model. Therefore, at this stage, the comparison operation does not traverse the entire silhouette information set  $S_2$  of node  $N_2$ . Rather, through new context added from the previous node  $N_1$ , only  $\ell_2$  comparisons are performed in the pertinent silhouette information part  $A_2$ , a subset of the entire set  $S_2$  at node  $N_2$  to yield  $V_2$ .

The best-case time complexity of comparison at this stage is again simply  $t$ , if the detected MO matches with silhouette information of  $N_2$  in its very first comparison.

The average-case time complexity is

$$\frac{1}{\ell_2} \sum_{i=1}^{\ell_2} t_i = \frac{\ell_2 + 1}{2} t, \tag{13}$$

where  $\ell_2$  is the last object of set  $A_2$ .

The worst-case time complexity of comparison is

$$t \sum_{i=1}^{\ell_2} 1 = \ell_2 t \tag{14}$$

if the detected MO matches with silhouette information of  $N_2$  in its very last comparison.

*Total Time Taken till the Second Node.* Concluding both nodes, the best-case time complexity is  $2t$ .

Similarly, the average-case time complexity is

$$\frac{m_1 + 1}{2} t + \frac{\ell_2 + 1}{2} t = \frac{m_1 + \ell_2 + 2}{2} t. \tag{15}$$

The worst-case time complexity is

$$m_1 t + \ell_2 t = (m_1 + \ell_2) t. \tag{16}$$

In rare cases, if the comparison operation yields incorrect result and a wrong object recognition message is sent to the second node, subsequent comparison to a subset  $A_i$  would yield a mismatch. Consequently, the second node once again has to traverse the entire silhouette information set  $S_2$  of node  $N_2$ .

Then, the worst-case time complexity in this situation is

$$\begin{aligned}
 t \sum_{i=1}^{m_1} 1 + t \sum_{i=1}^{\ell_2} 1 + t \sum_{i=1}^{m_2} 1 &= m_1 t + \ell_2 t + m_2 t \\
 &= (m_1 + \ell_2 + m_2) t.
 \end{aligned} \tag{17}$$

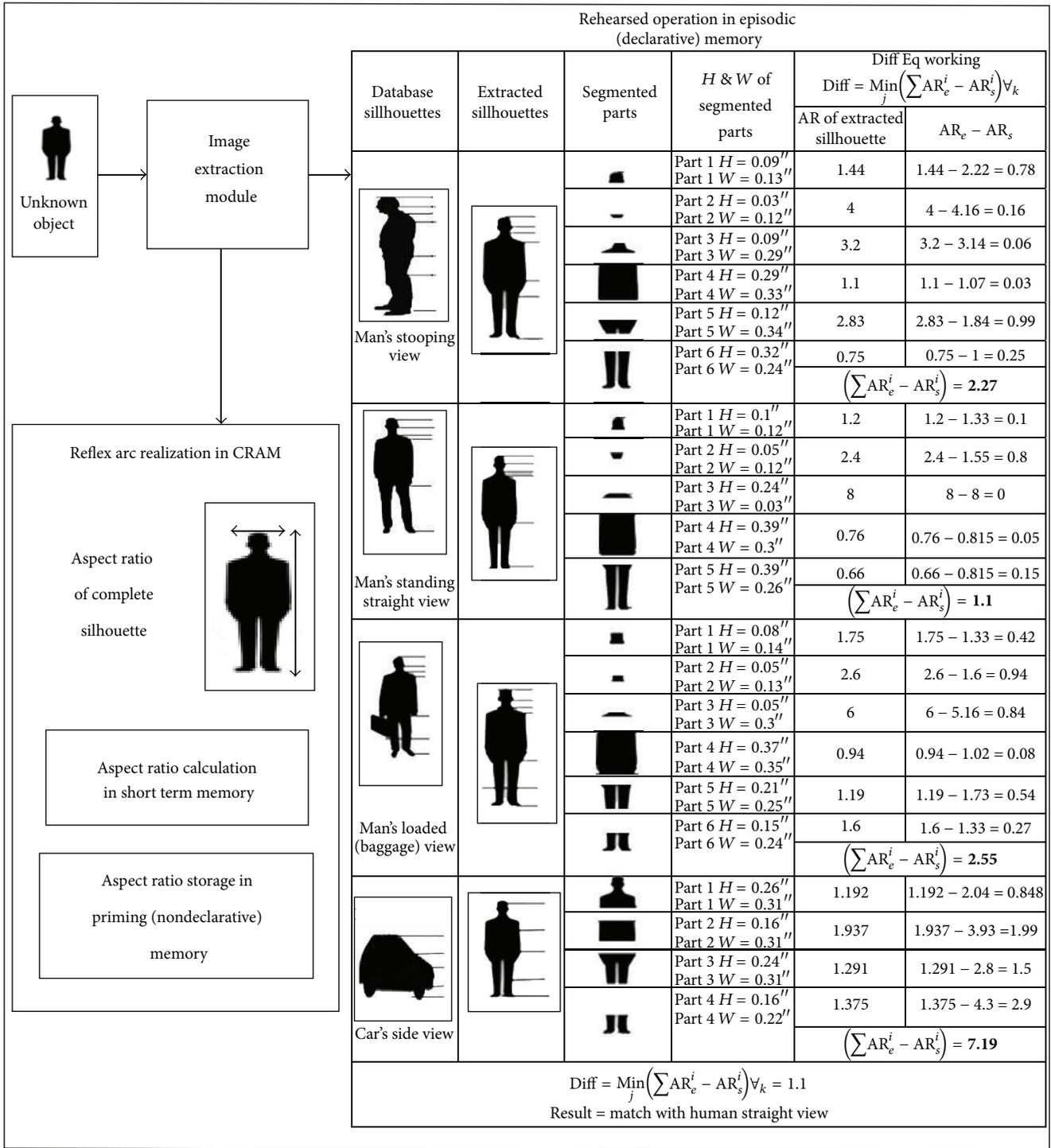


FIGURE 13: Function of episodic memory and realization of reflex through priming memory.

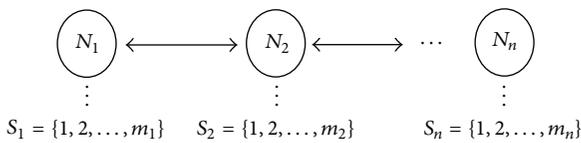


FIGURE 14: Trajectory of the detected MO.

Time Taken till nth Node. Similarly, for traversing  $n$  number of nodes' surety level as  $\sum V_i \geq V_{\text{Threshold}}$ , the time complexity of our model can be deduced through induction as follows.

TABLE 17: CRAM reflex response.

Captured image	Image	Elapsed time		Recognized image	
		VISTA	CRAM	VISTA	CRAM
Human standing		16.87 sec	2.02 sec	Human standing	Human standing
Car front view		24.40 sec	2.05 sec	Car front view	Car front view
Car side view		25.20 sec	2.28 sec	Car front view	Car front view

Best-case time complexity is  $nt$  if the defined silhouette information of all  $N_i$  is matched in  $S_1$  and all  $A_i$  ( $i$  iterates from 2 to  $n$ ) in their very first comparisons.

Average-case time complexity is

$$\frac{m_1 + 1}{2}t + \frac{\ell_2 + 1}{2}t + \dots + \frac{\ell_n + 1}{2}t$$

$$= \frac{1}{2} \left( m_1 + n + \sum_{a=\ell_2}^{\ell_n} a \right) t. \quad (18)$$

Worst-case time is

$$t \sum_{i=1}^{m_1} 1 + t \sum_{j=1}^{\ell_2} 1 + \dots + t \sum_{k=1}^{\ell_n} 1$$

$$= m_1 t + (\ell_2 + \ell_3 + \dots + \ell_n) t = \left( m_1 + \sum_{a=\ell_2}^{\ell_n} a \right) t. \quad (19)$$

In very rare cases, if the comparison operations yield incorrect results and wrong object recognition messages are sent to every node, subsequent comparisons to pertaining subsets  $A_i$  would yield mismatches. Consequently, every node once again has to traverse the entire silhouette information set  $S_i$  of node  $N_i$ .

Then, the worst-case time complexity in this situation will be

$$t \sum_{i=1}^{m_1} 1 + t \sum_{i=1}^{\ell_2} 1 + t \sum_{i=1}^{m_2} 1 + t \sum_{i=1}^{\ell_3} 1 + t \sum_{i=1}^{m_3} 1 + \dots + t \sum_{i=1}^{\ell_n} 1$$

$$+ t \sum_{i=1}^{m_n} 1 = m_1 t + m_2 t + m_3 t + \dots + m_n t + \ell_2 t + \ell_3 t$$

$$+ \dots + \ell_n t = (m_1 + m_2 + \dots + m_n) t$$

$$+ (\ell_2 + \ell_3 + \dots + \ell_n) t = t \sum_{a=m_1}^{m_n} a + t \sum_{a=\ell_2}^{\ell_n} a. \quad (20)$$

Equations (18) and (19) express the total time consumed by a context added system in average and worst environment, respectively, while (20) shows the time consumed by the system in a very rare case. However, in such a situation, the system will not generate a valid surety value. As we will see, it is almost the result in case of a contemporary context aware system under normal circumstances.

For a contemporary context aware system in average-case circumstances, the time complexity is

TABLE 18: CRAM accuracy.

Case number	MO aspect	Image	Image recognized?	
			VISTA	CRAM
1	Car front view		✓	✓
2	Car front view		✓	✗ (Object identified as car back side view)
3	Car left side view		✓	✓
4	Car left side view		✗ Object identified as car right side	✗ Object identified as car right side view
5	Human straight view (good light conditions)		✓	✓
6	Human straight view (good light conditions)		✓	✓
7	Human straight view (good light conditions)		✓	✓
8	Human straight view (bad light conditions)		✓	✓
9	Human straight view (bad light conditions)		✗ Object identified as human loaded view	✓

TABLE 18: Continued.

Case number	MO aspect	Image	Image recognized?	
			VISTA	CRAM
10	Human straight view (bad light conditions)		✓	✓

$$\begin{aligned} & \frac{m_1 + 1}{2}t + \frac{m_2 + 1}{2}t + \dots + \frac{m_n + 1}{2}t \\ &= \frac{1}{2} \left( n + \sum_{a=m_1}^{m_n} a \right) t. \end{aligned} \quad (21)$$

For a contemporary context aware system in worst-case circumstances, the time complexity is

$$\begin{aligned} & t \sum_{i=1}^{m_1} 1 + t \sum_{i=1}^{m_2} 1 + \dots + t \sum_{i=1}^{m_n} 1 = (m_1 + m_2 + \dots + m_n) t \\ &= \left( \sum_{a=m_1}^{m_n} a \right) t, \end{aligned} \quad (22)$$

$$\frac{1}{2} \left( m_1 + n + \sum_{a=\ell_2}^{\ell_n} a \right) t < \frac{1}{2} \left( n + \sum_{a=m_1}^{m_n} a \right) t, \quad (23)$$

$$\left( m_1 + \sum_{a=\ell_2}^{\ell_n} a \right) t < \left( \sum_{a=m_1}^{m_n} a \right) t. \quad (24)$$

Equation (23) depicts the comparison of (18) and (21). Similarly, (24) depicts the comparison of (19) and (22). It is evident that the total time consumed by the context added system comprising  $n$  nodes to track and recognize a MO is much less than that of a context aware system.

**6.2. Context Addition Is Recursive.** In order to assess the behavior of context addition, we consider the trajectory of MO through the same set  $N$  of  $n$  nodes as discussed in previous section. Suppose that, at node  $N_1$ , the percentage surety level of the detected MO is calculated as a value  $V_1$ . At node  $N_2$ , its local surety value  $V_2$  is calculated by adding to it the received surety value  $V_1$  of node  $N_1$ . The context addition of our architecture proceeds in this recursive way. The recurrence equation for  $V_n$  received from  $n$  nodes is as the following:

$$V_n = \begin{cases} V_1, & \text{for } n = 1, \\ V_n + V_{n-1}, & \text{for } n \geq 2. \end{cases} \quad (25)$$

## 7. Conclusions and Future Work

In this paper, we present context aware systems that evolve into autonomic, intelligent processing systems through the

incorporation of context addition. We have presented conceptual, architectural, and deployment aspects of context added system. Through establishing an analogy between context added system and human anatomy of memory, we have proposed the incorporation of reflex arc into context aware systems. We have demonstrated that both context awareness and reflex arc can be embedded into visual context aware systems through prototype implementation.

As part of our future work, the recursive behavior of the CARM will be analyzed more specifically to calculate and compare the average performance variance of context added versus context aware systems and a possible trade-off between latency, precision, and autonomicity. Finally, the authors are quite optimistic to study and analyze the self-learning and reflex response features of the proposed model in the domain of Internet of Things to contribute towards the initiative of *Intelligent Civilization*.

## Competing Interests

The authors declare that they have no competing interests.

## References

- [1] S. Fayed, S. M. Youssef, A. El-Helw, M. Patwary, and M. Moniri, "Adaptive compressive sensing for target tracking within wireless visual sensor networks-based surveillance applications," *Multimedia Tools and Applications*, 2015.
- [2] S. Jabbar, A. H. Akbar, S. Zafar, M. M. Quddoos, and M. Hussain, "VISTA: achieving cumulative VISION through energy efficient Silhouette recognition of mobile Targets through collaboration of visual sensor nodes," *EURASIP Journal on Image and Video Processing*, vol. 2014, article 32, 24 pages, 2014.
- [3] M. Baldauf, S. Dustdar, and F. Rosenberg, "A survey on context-aware systems," *International Journal of Ad Hoc and Ubiquitous Computing*, vol. 2, no. 4, pp. 263–277, 2007.
- [4] N. Ryan, J. Pascoe, and D. Morse, "Enhanced reality fieldwork: the context aware archaeological assistant," vol. 750 of *Bar International Series*, pp. 269–274, 1999.
- [5] G. D. Abowd, C. G. Atkeson, J. Hong, S. Long, R. Kooper, and M. Pinkerton, "Cyber guide: a mobile context-aware tour guide," *Wireless Networks*, vol. 3, no. 5, pp. 421–433, 1997.
- [6] K. Cheverst, N. Davies, K. Mitchell, A. Friday, and C. Efstathiou, "Developing a context-aware electronic tourist guide: some issues and experiences," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 17–24, ACM, April 2000.

- [7] A. K. Dey, "Context-aware computing: the Cyber Desk project," in *Proceedings of the AAAI Spring Symposium on Intelligent Environments*, pp. 51–54, Menlo Park, Calif, USA, March 1998.
- [8] P. J. Brown, "The stick-e document: a framework for creating context-aware applications," in *Proceedings of the Electronic Publishing (EP '96)*, pp. 182–196, Palo Alto, Calif, USA, January 1996.
- [9] G. D. Abowd, A. K. Dey, P. J. Brown, N. Davies, M. Smith, and P. Steggles, "Towards a better understanding of context and context-awareness," in *Handheld and Ubiquitous Computing*, vol. 1707 of *Lecture Notes in Computer Science*, pp. 304–307, Springer, Berlin, Germany, 1999.
- [10] R. Hull, P. Neaves, and J. Bedford-Roberts, "Towards situated computing," in *Proceedings of the 1st International Symposium on Wearable Computers. Digest of Papers*, pp. 146–153, IEEE, October 1997.
- [11] A. Zimmermann, M. Specht, and A. Lorenz, "Personalization and context management," *User Modelling and User-Adapted Interaction*, vol. 15, no. 3–4, pp. 275–302, 2005.
- [12] S. Loke, *Context-Aware Pervasive Systems: Architectures for a New Breed of Applications*, CRC Press, New York, NY, USA, 2006.
- [13] N. Baker, M. Zafar, B. Moltchanov, and M. Knappmeyer, "Context-aware systems and implications for future internet," in *Towards the Future Internet*, G. Tselentis, J. Domingue, A. Galis et al., Eds., pp. 335–344, IOS Press, Amsterdam, The Netherlands, 2009.
- [14] A. Schmidt, "Context-aware computing: context-awareness, context-aware user interfaces, and implicit interaction," in *The Encyclopedia of Human-Computer Interaction*, The Interaction Design Foundation, 2nd edition, 2013.
- [15] S. Chatterjea, T. Nieberg, N. Meratnia, and P. Havinga, "A distributed and self-organizing scheduling algorithm for energy-efficient data aggregation in wireless sensor networks," *ACM Transactions on Sensor Networks*, vol. 4, no. 4, article 20, 2008.
- [16] S. Jin and K. Li, "LBCS: a load balanced clustering scheme in wireless sensor networks," in *Proceedings of the 3rd International Conference on Multimedia and Ubiquitous Engineering (MUE '09)*, pp. 221–225, Qingdao, China, June 2009.
- [17] A. Zafeiropoulos, P. Gouvas, A. Liakopoulos, G. Mentzas, and N. Mitrou, "NEURON: enabling autonomy in wireless sensor networks," *Sensors*, vol. 10, no. 5, pp. 5233–5262, 2010.
- [18] R. Žontar, M. Heričko, and I. Rozman, "Taxonomy of context-aware systems," *Elektrotehniški Vestnik*, vol. 79, no. 1–2, pp. 41–46, 2012.
- [19] C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos, "Context aware computing for the internet of things: a survey," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 414–454, 2014.
- [20] A. Bikakis, T. Patkos, G. Antoniou, and D. Plexousakis, "A survey of semantics-based approaches for context reasoning in ambient intelligence," in *Constructing Ambient Intelligence*, vol. 11 of *Communications in Computer and Information Science*, pp. 14–23, Springer, Berlin, Germany, 2008.
- [21] C. Venezia and C. A. Licciardi, "Improve ubiquitous web applications with context awareness," in *Proceedings of the 11th International Conference on Intelligence in Service Delivery Networks (ICIN '07)*, vol. 16, pp. 649–656, October 2007.
- [22] B. Moltchanov, M. Knappmeyer, C. A. Licciardi, and N. Baker, "Context-aware content sharing and casting," in *Proceedings of ICIN*, Bordeaux, France, October 2008.
- [23] H. Chen, T. Finin, and A. Joshi, "An intelligent broker for context-aware systems," in *Proceedings of the 5th International Conference on Ubiquitous Computing (UbiComp '03)*, vol. 3, pp. 183–184, Seattle, Wash, USA, October 2003.
- [24] 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.
- [25] P. Floréen, M. Przybilski, P. Nurmi et al., "Towards a context management framework for MobiLife," in *Proceedings of the 14th IST Mobile & Wireless Summit*, pp. 20–28, June 2005.
- [26] C. Cordier, F. Carrez, H. Van Kranenburg et al., "Addressing the challenges of Beyond 3G service delivery: the SPICE service platform," in *Proceedings of the Workshop on Applications and Services in Wireless Networks (ASWN '06)*, Berlin, Germany, May 2006.
- [27] J. C. Yelmo, J. M. del Álamo, R. Trapero et al., "A user-centric service creation approach for Next Generation Networks," in *Proceedings of the 1st ITU-T Kaleidoscope Academic Conference Innovations in NGN: Future Network and Services (K-INGN '08)*, pp. 211–218, IEEE, Geneva, Switzerland, May 2008.
- [28] L. Lamorte, C. A. Licciardi, M. Marengo et al., "A platform for enabling context aware telecommunication services," in *Proceedings of the 3rd Workshop on Context Awareness for Proactive Systems*, Guildford, UK, June 2007.
- [29] E. S. Barrenechea, P. S. Alencar, R. Blanco, and D. Cowan, "Context-awareness and adaptation in distributed event-based systems," Tech. Rep. CS-2011-14, David R. Cheriton School of Computer Science, University of Waterloo, Waterloo, Canada, 2011.
- [30] J. Kane, B. Tang, Z. Chen et al., "Reflex-tree: a biologically inspired parallel architecture for future smart cities," in *Proceedings of the 44th International Conference on Parallel Processing (ICPP '15)*, pp. 360–369, Beijing, China, September 2015.
- [31] P. K. Anokhin, *Biology and Neurophysiology of the Conditioned Reflex and Its Role in Adaptive Behavior*, vol. 3 of *International Series of Monographs in Cerebrovisceral and Behavioral Physiology and Conditioned Reflexes*, Elsevier, Philadelphia, Pa, USA, 2013.
- [32] I. P. Pavlov and G. V. Anrep, *Conditioned Reflexes*, vol. 614, Courier Corporation, 2003.
- [33] J. F. Iles, M. Stokes, and A. Young, "Reflex actions of knee joint afferents during contraction of the human quadriceps," *Clinical Physiology*, vol. 10, no. 5, pp. 489–500, 1990.
- [34] E. Onur, C. Ersoy, and H. Deliç, "How many sensors for an acceptable breach detection probability?" *Computer Communications*, vol. 29, no. 2, pp. 173–182, 2006.
- [35] Y.-X. Li, H.-S. Shi, and S.-P. Zhang, "An energy-efficient MAC protocol for wireless sensor network," in *Proceedings of the 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE '10)*, pp. V4619–V4623, IEEE, Chengdu, China, August 2010.
- [36] C. Ford and I. Stange, "A framework for generalizing public safety video applications to determine quality requirements," in *Proceedings of the IEEE Conference on Multimedia Communications, Services & Security*, Krakow, Poland, May 2010.
- [37] A. K. Bashir, A. H. Akbar, S. A. Chaudhary, C. S. Hussain, and K.-H. Kim, "Collaborative detection and agreement protocol for routing malfunctioning in wireless sensor networks," in *Proceedings of the 8th International Conference Advanced Communication Technology (ICACT '06)*, vol. 1, pp. 327–332, IEEE, Phoenix Park, Ireland, February 2006.

- [38] S. Chandra, C. S. Ellis, and A. Vahdat, "Managing the storage and battery resources in an image capture device (digital camera) using dynamic transcoding," in *Proceedings of the 3rd ACM International Workshop on Wireless Mobile Multimedia*, pp. 73–82, ACM, August 2000.
- [39] L. Xu, S. Zhang, Z. He, and Y. Guo, "The comparative study of three methods of remote sensing image change detection," in *Proceedings of the IEEE 17th International Conference on Geo Informatics*, pp. 1–4, 2009.
- [40] M. T. Ullman, "Contributions of memory circuits to language: the declarative/procedural model," *Cognition*, vol. 92, no. 1-2, pp. 231–270, 2004.
- [41] E. Tulving, "Episodic and semantic memory 1," *Organization of Memory*, vol. 381, no. 4, article e402, 1972.
- [42] D. L. Schacter, D. T. Gilbert, and D. M. Wegner, "Semantic and episodic memory," in *Psychology*, pp. 240–241, 2nd edition, 2011.
- [43] M. Hilker, J. Schwachtje, M. Baier et al., "Priming and memory of stress responses in organisms lacking a nervous system," *Biological Reviews*, 2015.
- [44] D. L. Schacter and E. Tulving, *Memory Systems 1994*, MIT Press, Cambridge, Mass, USA, 1994.

## Research Article

# Development and Coverage Evaluation of ZigBee-Based Wireless Network Applications

Fei Ding<sup>1,2</sup> and Aiguo Song<sup>3</sup>

<sup>1</sup>*School of Information Science and Engineering, Southeast University, Nanjing 210096, China*

<sup>2</sup>*R&D Center, China Mobile Group Jiangsu Co., Ltd., Nanjing 210029, China*

<sup>3</sup>*School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China*

Correspondence should be addressed to Aiguo Song; [a.g.song@seu.edu.cn](mailto:a.g.song@seu.edu.cn)

Received 7 December 2015; Revised 12 February 2016; Accepted 18 February 2016

Academic Editor: Gwanggil Jeon

Copyright © 2016 F. Ding and A. Song. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Network coverage is one of the basic issues for information collection and data processing in ZigBee-based wireless sensor networks. Each node may be randomly distributed in a monitoring area, reflecting the network event of tracking in ZigBee network applications. This paper presents the development and coverage evaluation of a ZigBee-based wireless network application. A stack structure node available for home service integration is proposed, and all data of sensing nodes with an adaptive weighted fusion (AWF) processing are passed to the gateway and through the gateway to reexecute packet processing and then reported to the monitoring center, which effectively optimize the wireless network to the scale of the data processing efficiency. The linear interpolation theory is used for background graphical user interface so as to evaluate the working status of each node and the whole network coverage case. A testbed has been created for validating the basic functions of the proposed ZigBee-based home network system. Network coverage capabilities were tested, and packet loss and energy saving of the proposed system in longtime wireless network monitoring tasks were also verified.

## 1. Introduction

With recent *wireless sensor network* (WSN) development, more and more sensors and actuators for monitoring and control are embedded with wireless communication standards, which can be connected together to form an autonomous network. Meanwhile, ZigBee-based technology is considered one of the major communication advances in recent years, since it offers the basis for mesh networking, full area coverage, and development of independent cooperative services and applications [1].

Extensive research is underway using this concept in different areas, such as regional environmental monitoring and management [2], animal presence and pasture time monitoring in an extended area [3], medication management and health care system [4], and, in particular, *home network application* (HNA). For example, the IoT potential for HNAs has been reported in [5–8]. An intelligent self-adjusting sensor for smart home services based on ZigBee

Communications is proposed in [5]. A smart home testbed based on the pedagogical model of project-based learning (PBL) for undergraduate education is proposed in [6]. A wireless home automation network for indoor surveillance is presented in [7]. References [8, 9] proposed smart and wireless home energy management scenario synchronously. These show potential identification capacities for self-configuration, comprehensive management, and communication capabilities.

However, there are still some challenges in designing ZigBee-based home network system and applications. Firstly, according to the current situation, the ZigBee-based service terminals need to provide the total solution to improve the integrated efficiency. Meanwhile, the ZigBee module will be used as a supplementary means of integration. Secondly, the ZigBee-based service node uses a limited power resource, such as a battery. Thus, the network lifetime is greatly influenced by the battery lifetime. Last but in no means least, compared with the ZigBee network itself, users are more

TABLE I: Fusion algorithms comparison.

Classification	Title	Remarks	
Classical fusion algorithm	Statistical method	Bias algorithm D-S theory	
	Estimation method	Maximum likelihood estimation	Need a priori knowledge and need to meet more set conditions
		Kalman filter	The integrity of the data protection is good and applicable for high robustness applications
		Least squares Data fusion	
New fusion algorithm	Information theory	Cluster analysis	No need for a priori conditions, but the computation is large and the energy consumption is high
	Artificial intelligence	Fuzzy estimation	
		Neural network	

concerned about the visual perception. For example, how to effectively monitor, model, and manage monitoring processes is a critical task for current situation.

In ZigBee wireless network, sensor nodes scale deployment and, with limited power, it is necessary to implement data fusion procedure and reduce data reporting frequency and improve the measure precision and meet the long-term, reliable coverage monitoring needs.

At present, the main fusion algorithms are shown in Table 1, and all kinds of algorithms have their own characteristics and application scenarios [10–15]. And for the energy constrained ZigBee node, the data fusion algorithm is a good choice, which mainly includes the *arithmetic mean method* (AMM), the batch estimation algorithm, and the *adaptive weighted fusion* (AWF) algorithm. Among them, the AMM is the most simple, but its fusion precision is low; batch estimation algorithm can avoid the negative impact brought by the abnormal data, but it can ensure that the measurement accuracy of the sensor is highly consistent; AWF algorithm is relatively simple to implement, of high precision, only relying on multisensor data measurements, and applicable to the ZigBee network node, respectively, to perform data acquisition and monitoring coverage of the scene.

In this study, we designed a *ZigBee-based intelligent self-adjusting node* (ZiSAN) for wireless network applications. The ZiSAN, using a compact stack type structure, provides a *microoperation system* (MOS) for ZigBee network management and service scheduling. Then, a development and coverage evaluation of ZigBee-based wireless network application is proposed. The remainder of this paper is organized as follows. Section 2 proposes the ZigBee-based networking architecture for environmental monitoring system. Section 3 describes the design method of the ZigBee-based sensing node, data processing model, workflow, and procedure. Section 4 illustrates the testbed setup and results and finally some conclusions are presented.

## 2. System Overview

As shown in Figure 1, a ZigBee-based WSN architecture adapted to home network application mainly includes three types of nodes: sink node, ZigBee router (ZR), and ZigBee end node (ZE), respectively. The sink node is the *ZigBee-based gateway* (ZBGW) which is also the coordinator (ZC) of the

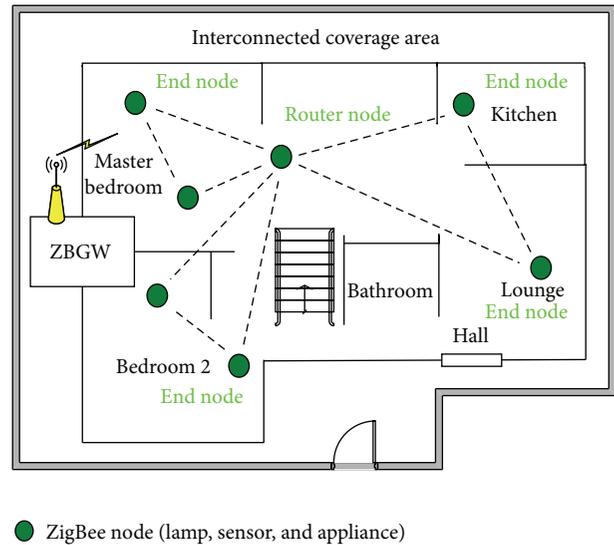


FIGURE 1: ZigBee-based network scheme for home applications.

ZigBee network, and it is responsible for reporting light data to user service side.

In ZigBee-based wireless monitoring area, two nodes that need to communicate may be beyond direct communication range, so the partial nodes are adopted to be configured as router nodes (ZR nodes) and then all the nodes working in the 2.4 frequency band and building up a ZigBee-based mesh network structure enable the end nodes (ZE nodes) to use the ZR nodes for data forwarding. In the mesh topology, a “route discovery” feature of each node is configured which allows the network to find the best available route for data transmission. Likewise, if there is a bad radio link (coverage hole due to the failure communication) between two nodes, that coverage hole can be overcome by rerouting around the area of bad service.

ZigBee nodes (ZE or ZR) integrated with the home security terminal (e.g., Doorsensor and smoke detector), environmental monitoring terminal (such as formaldehyde or PM2.5 monitor), and controller terminal (such as remote control device or a socket), and all of them build a mesh structure network. The sensing node receives the control instruction through the ZBGW and carries out the real-time

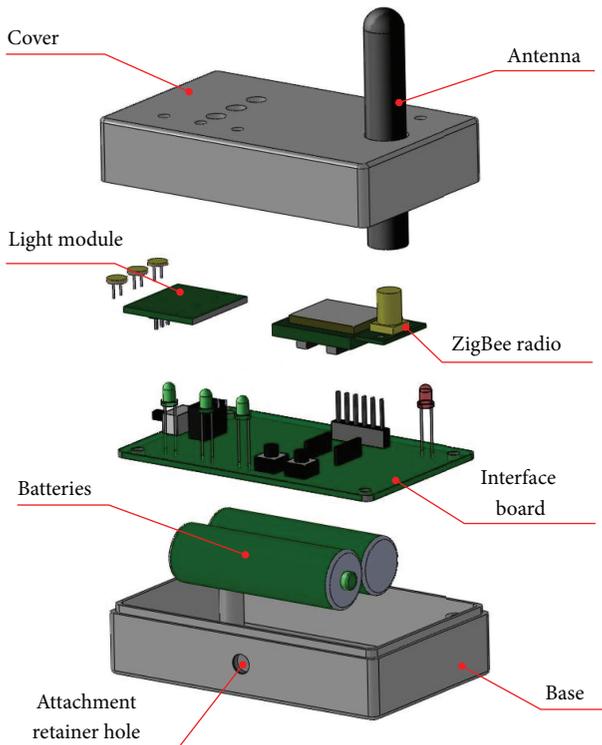


FIGURE 2: Exploded view of the proposed ZiSAN stack structure.

monitoring task. Node failure (the power supply is under pressure, being destroyed, etc.) means the coverage with a hole and the node in the failure area to report to an alarm automatically. According to the application requirement of the ZigBee wireless network, the packet loss retransmission mechanism (no more than 3 times) will be executed when the communication between the nodes appeared, so as to ensure the effective coverage of the wireless network application system.

### 3. Design Methods

**3.1. ZiSAN Structure Decomposition.** In a home network system, the ZiSAN mainly acts as a ZR or ZE. In order to facilitate the integrated development of home terminal and ZigBee technology, the ZiSAN uses the stack structure convenient for sensor replacement and expansion. Figure 2 shows the structure of the proposed ZiSAN used for light monitoring in an exploded view. The whole architecture includes a ZigBee radio module, a light sensing module, an interface board, a case plastic, and so forth. The ZiSANS powered by 2 AA batteries or *direct current* (DC) are integrated on other monitoring or control devices and deployed at different places of the home monitoring area.

The ZigBee radio is a microembedded system, which integrated with a *microcontrol unit* (MCU) and *microoperation system* (MOS) software, and it connected with the interface board through a 22-pin interface, so as to output its *general purpose input output* (GPIO) and *cluster communication port*

(COM) to the interface board. The ZigBee radio can be configured with 4 different power levels (8.25 mW–66 mW) and it is also equipped with an onboard antenna or a cylindrical antenna, so as to make it possible for the user to fit all kinds of monitoring field. An arrangement structure of 3 light sensors enables the ZiSAN to distinguish between light and darkness, to read the light intensity in wireless network applications.

**3.2. Data Processing Strategy.** A ZigBee-based wireless network scenario exhibits a ZBGW and a number of ZiSANS; the ZiSANS implement coverage monitoring task and cause a large scale of homogeneous data; such data are gathered by the ZBGW and then reported to the monitoring center. If ZBGW collected each piece of sensing data of the sensing node, without any processing and directly reporting to the monitoring center, in the ZigBee wireless channel, congestion of ZBGW serial data is easy to be caused. Therefore, it is necessary for ZBGW to reexecute package processing to be reported again, so as to optimize the efficiency of the transmission channel.

The data processing workflow is shown in Figure 3, and the specific evaluation processes are as follows.

- (S1) Each sensing node is installed with multiple light sensors, and continuous sampling is performed simultaneously.
- (S2) The multisensor sampling data is processed by the AWF model for data fusion.

The AWF process of executing includes the following steps: firstly, the sampling values of each sensor are  $x_i$  ( $i = 1, 2, \dots, n$ ), and they are independent of each other, so as to evaluate the variance of the sensing node; then, according to the multivariate function theory, the weighted factor  $\omega_i$  ( $i = 1, 2, \dots, n$ ) can be obtained by the value corresponding to the minimum *total mean square error* (TMSE) [13–15]; secondly, the sensing data are fused with the WAF model, thereby obtaining the final real-time fusion value of the sensor node.

- (S3) The same type sensor data is packed into a stack queue through the ZBGW and reexecution of packet processing within set time threshold is performed, with the whole packet length being less than the set threshold.
- (S4) If Timer 1 overflows, Step 3 is implemented; otherwise Step 5 is performed.
- (S5) The ZBGW reports data (regardless of whether the message packet meets the full packet condition) to the monitoring center.
- (S6) The monitoring center displays sensor data in real-time and positions monitoring events.
- (S7) Check the data queue of the ZBGW if there exists a full message packet.
- (S8) If there is a full message packet that complies with the requirements, then go to (S5); otherwise (S7) is reexecuted.

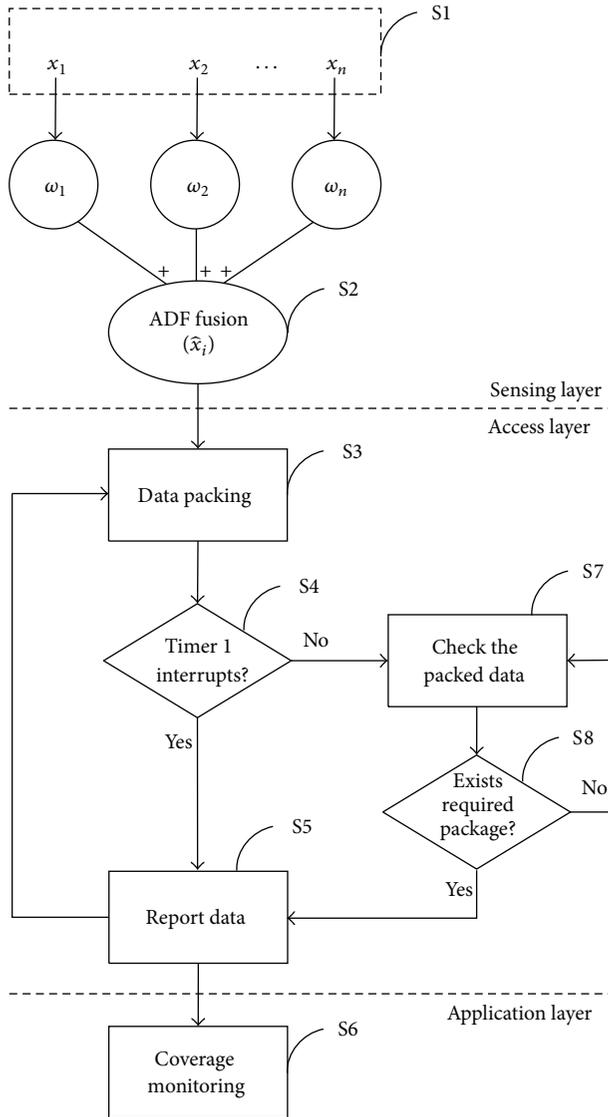


FIGURE 3: Data processing and reporting workflow.

**3.3. Event Monitoring Model.** According to the ZiSAN data transfer to the user client server through the ZBGW, the linear interpolation principle is adopted to obtain a full rendering. As shown in Figure 4,  $2 * 2$  pixels represent the wireless monitoring area converging with 4 nodes, and then it can be turned into a 9-pixel picture through interpolation, which reflects that the monitoring area of each ZiSAN is smaller, but the behavior monitor of each node can be evaluated more accurately. Hence, through the background *graphical user interface* (GUI), users can directly monitor the working status of the wireless network and timely process the monitoring events.

The capital letters refer to the original pixel, and lowercase letters refer to the new pixel obtained through interpolation; define its linear interpolation as

$$\begin{aligned}
 ab &= \frac{A+B}{2}, \\
 cd &= \frac{C+D}{2}, \\
 ac &= \frac{A+C}{2}, \\
 bd &= \frac{B+D}{2}, \\
 abcd &= \frac{ab+cd}{2} = \frac{A+B+C+D}{4}.
 \end{aligned} \tag{1}$$

The monitoring and analysis of the wireless network can be transformed into interpolation evaluation with multiple ZiSANS and deduce

$$\begin{aligned}
 ab &= A + \frac{B-A}{2}, \\
 cd &= C + \frac{D-C}{2} \\
 &\vdots
 \end{aligned} \tag{2}$$

The parameter of  $ab$  and  $cd$  can be evaluated from the vertical direction. Similarly, the parameter of  $ac$  and  $bd$  can be evaluated from the horizontal direction. Finally we obtain  $abcd$ .

The evaluation process will be illustrated as follows. The coordinate of pixel C is defined as  $(x_1, y_1)$ , and its real-time light data is  $L_1$ ; with the same principle, the coordinate of pixel A is  $(x_1, y_2)$ ; its real-time light data is  $L_2$ . The coordinate of pixel D is  $(x_2, y_1)$ ; its real-time light data is  $L_3$ ; the coordinate of pixel B is  $(x_2, y_2)$ ; its real-time light data is  $L_4$ .

Based on the light intensity data,  $x, y$  coordinates, (3) is obtained as follows:

$$L(x, y) = a_0 + a_1x + a_2y + a_3xy, \tag{3}$$

where  $x_1 = K \cdot N$ ,  $x_2 = (K + 1) \cdot N$ ,  $y_1 = J \cdot N$ , and  $y_2 = (J + 1) \cdot N$ .

Then, it comes to an equation set

$$\begin{aligned}
 a_0 + a_1x_1 + a_2y_1 + a_3x_1y_1 &= L_1, \\
 a_0 + a_1x_1 + a_2y_2 + a_3x_1y_2 &= L_2, \\
 a_0 + a_1x_2 + a_2y_1 + a_3x_2y_1 &= L_3, \\
 a_0 + a_1x_2 + a_2y_2 + a_3x_2y_2 &= L_4.
 \end{aligned} \tag{4}$$

The equation is obtained as follows:

$$\begin{bmatrix} 1 & x_1 & y_1 & x_1y_1 \\ 1 & x_1 & y_2 & x_1y_2 \\ 1 & x_2 & y_1 & x_2y_1 \\ 1 & x_2 & y_2 & x_2y_2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} L_1 \\ L_2 \\ L_3 \\ L_4 \end{bmatrix}. \tag{5}$$

Therefore, (5) is solved for  $a_0, a_1, a_2$ , and  $a_3$  and substituted in

$$Z(x, y) = a_0 + a_1x + a_2y + a_3xy. \tag{6}$$

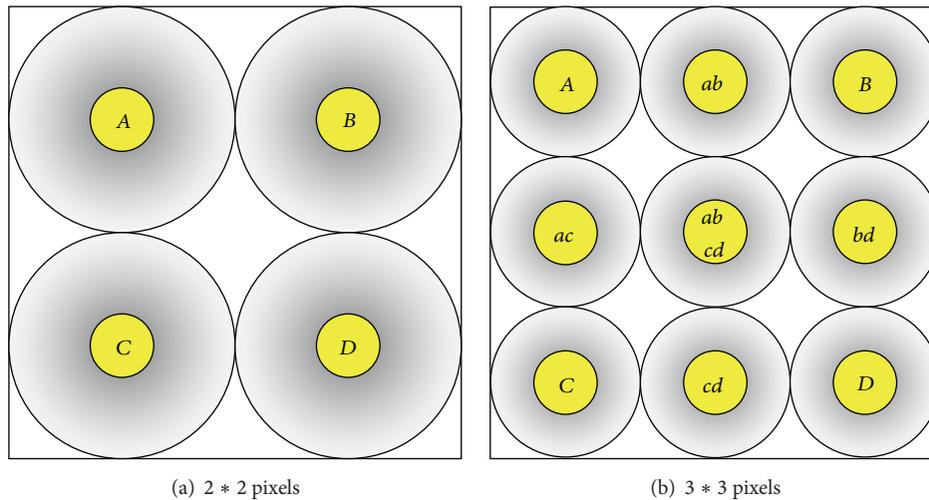


FIGURE 4: Light rendering with linear interpolation.

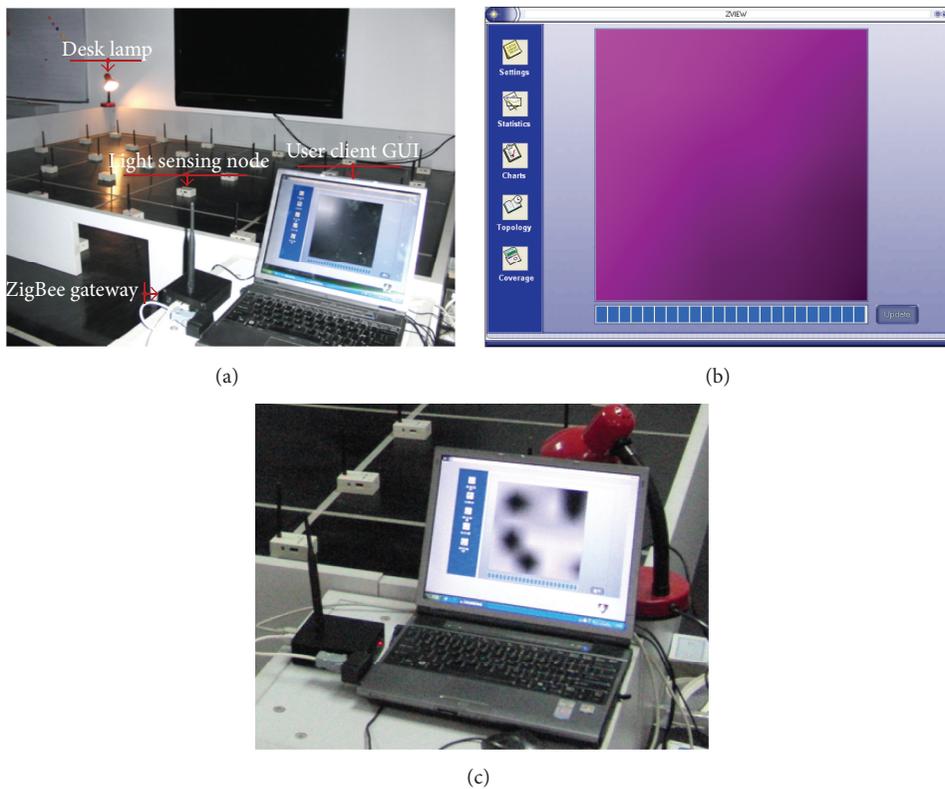


FIGURE 5: Environmental coverage experiments on the ZigBee-based wireless network: (a) testbed setup for the demonstration experiment; (b) graphic user interfaces running on PC; (c) nodes failure during lower voltage.

## 4. Experiments

**4.1. Prototype System.** A testbed of the proposed ZigBee-based home network system has been created for demonstration. The testbed setup is shown in Figure 5 which consists of a ZigBee gateway and 36 ZiSANS for mesh networking, and the schematic diagram of the testbed is shown in Figure 6.

The ZiSANS are deployed regularly in a  $250 \text{ cm} \times 250 \text{ cm}$  monitoring area, working on the minimum power level. The ZBGW is connected with the monitoring GUI through a serial port of PC. The ZiSAN senses the light intensity and transmits it to the ZBGW synchronously and finally reports to the background GUI. The user through the GUI can conveniently evaluate the light intensity of each monitoring point and evaluate the light distribution trend of the

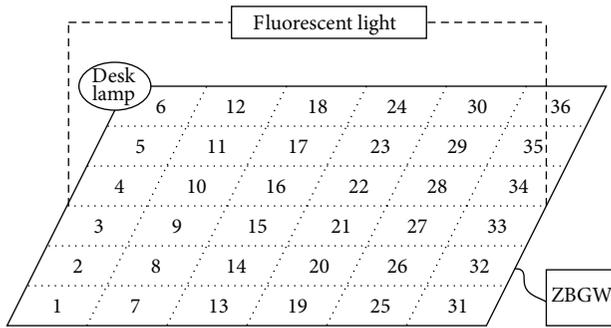


FIGURE 6: Schematic diagram showing the process of the coverage test.

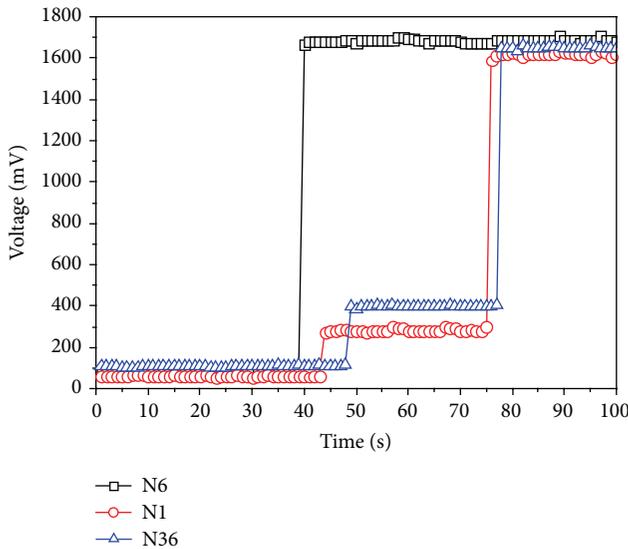


FIGURE 7: Sample results of the indoor light experiments.

whole proposed ZigBee-based wireless network monitoring scheme. If a ZiSAN fails (under voltage alarm, network failure, etc.), the user can visually monitor the network events.

## 5. Experimental Results

Figure 7 shows the real-time sample results of the ZiSANS (the adopted node number is 1, 6, and 36, resp.). First the ceiling and the desk lamp are both closed and the real-time light intensities of the 3 nodes are all very small. Then, the desk lamp located in the light sensing ZiSAN 6 is turned on, and node 6 has the maximum uplift amplitude of the real-time light data curve. Finally, the indoor ceiling lamp is switched on, and all of these 3 nodes have a larger uplift.

Figure 8 shows the light gradient changes of the monitoring testbed. According to the 3D curves, it is easy to know the illumination distribution of the test area. Meanwhile, the actual measurement shows that even though the ZiSANS in the monitoring area are arranged tightly, they can access the network once the ZiSANS start. And the real-time light sensing data will be updated every 2 seconds.

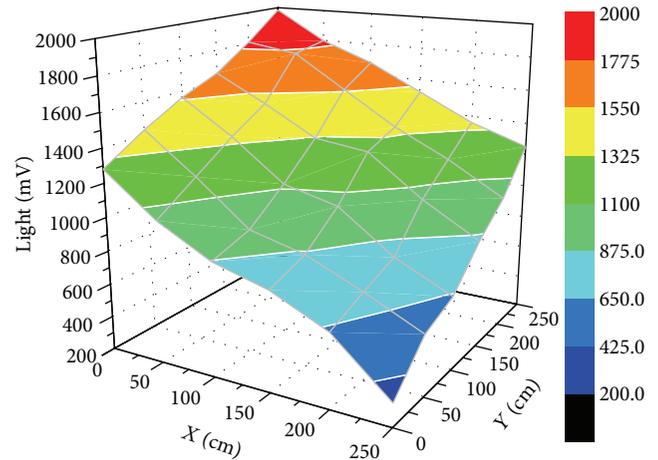


FIGURE 8: Light gradient measured by the light sensing nodes during the coverage test.

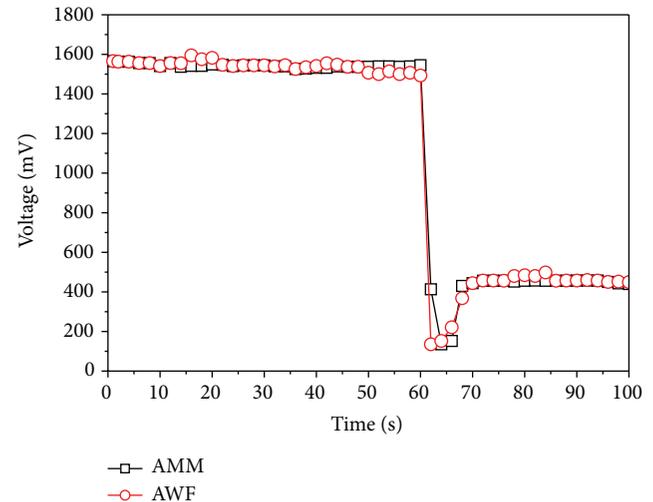


FIGURE 9: Data fusion comparison results.

Each ZiSAN mounted 3 light sensors and executed data acquisition once per second and continuously executed data acquisition 2 times. The light sensing data is processed by the AWF model. Node number 23 is selected for light real-time data collection and fusion, and the test result is shown in Figure 9. The first 60 seconds of the light curve is influenced by an indoor fluorescent light; the fluorescent light is turned off at 60 seconds and then the desk lamp is turned on nearby node number 6, followed by the curve of the increasing trend (amplitude is relatively small). As shown in Figure 9, when the light measured data processing by the AWF and the AMM (in most cases the fusion results remain basically the same) can be mutually replaced, only near the 15 seconds, 40 seconds, 50–70 seconds, 80 seconds, there exists some dynamic change, and the main reason is that the AWF model due to the weights can be automatically updated to the measured values, can be more sensitive to light, and has a better performance than that of the arithmetic mean method.

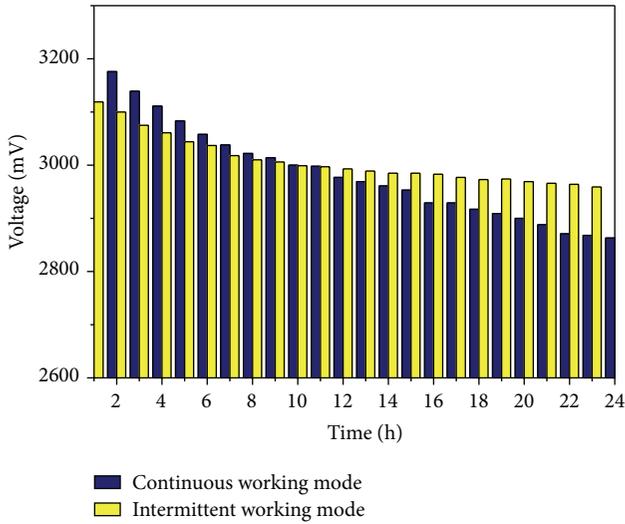


FIGURE 10: Node power consumption comparison results with different working mode.

Figure 10 is the power consumption comparison results of the ZiSAN in different operating modes. Continuous work mode is to perform data collection and data reporting every second; the intermittent mode of work is to collect data per second, but every two seconds after the data collection through AWF processing and then report, and it shows that the intermittent operation mode can optimize the power consumption of the node.

For ZBGW, continuous work mode is that real-time light data of the ZiSAN per second is received and reported to the monitoring center; reexecuting package processing mode refers to receiving the light data of the ZiSAN every 2 seconds and implementing data fusion of multinode's data and then transmitting to avoid the parallel wireless transmission and serial event triggered communication per second. Figure 11 shows the power consumption comparison results of the different working mode; when each node sends real-time sampling data per second, the power supply of the ZBGW is close to the configured threshold (2.5 V) after 24 hours and that of the reexecuting packet processing mode is better improved.

Some packet loss tests were carried out to evaluate the network stability of our ZigBee-based wireless monitoring system. ZiSANS are collected and reported one light dataset per second. A 16-byte light data protocol of the system which consists of packet header, MAC (medium access control), NetID, SenData, volt, and CRC (*Cyclical Redundancy Check*) is transferred every second and received by the PC GUI. The packet header is 0xAA and 0x55, and the MAC address, network address, real-time data, remaining energy voltage, and CRC byte are the corresponding components of the ZiSAN. As shown in Figure 12, packet loss test was performed 10 times to obtain the average of the failure probabilities; the right color represents the average packet loss number in 1000 seconds. The ZigBee multiple RF (radio frequency) communication links vary over location due to their strong

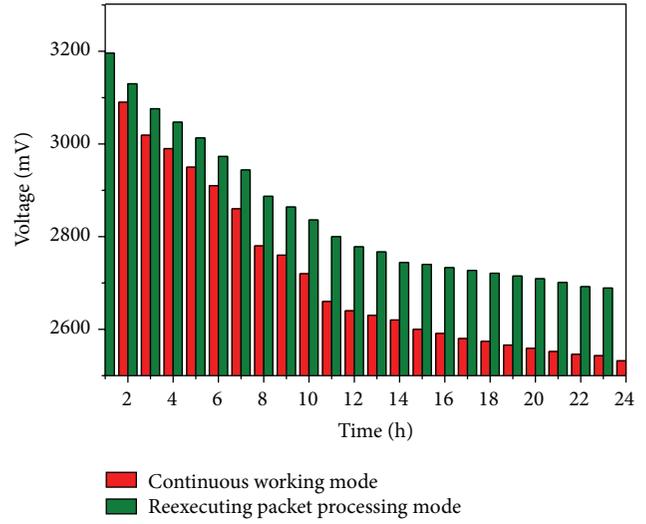


FIGURE 11: Analysis with power management of the ZBGW.

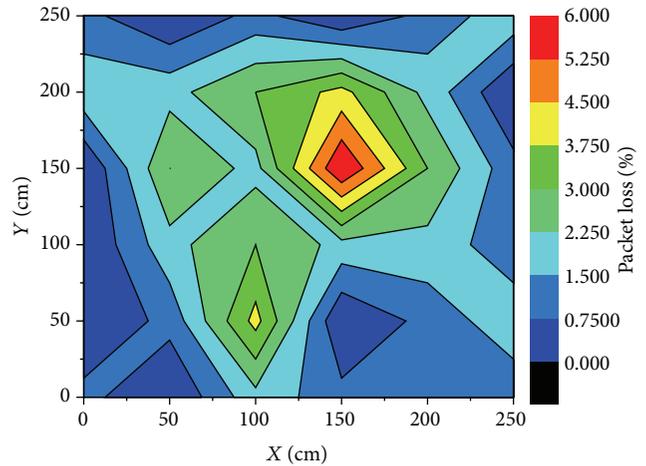


FIGURE 12: Packet loss evaluated by the ZigBee-based networking nodes.

correlation to the physical environment. The packet loss of the ZiSAN deployed in the middle area is higher than that of the edge node and the retransmit mechanism is configured in the application layer so as to improve the wireless communication.

Reexecuting packet processing method can effectively improve the packet loss of the ZigBee wireless network, so as to optimize the monitoring center of the screen refresh rate (e.g., a node data loss will display a coverage hole on the monitoring GUI). The packet loss test results of node number 23 in different working mode are shown in Figure 13. The initial stage of ZigBee network packet loss rate is higher. If the sampling data of the ZiSAN are not processed and fused by the ZBGW, the minimum refresh time is 1 second and the packet loss rate is about 3.5%. In addition, when packet loss occurs, the retransmission mechanism will consume extra energy. In contrast, when the reexecuting packet processing mechanism is adopted, the refresh time is

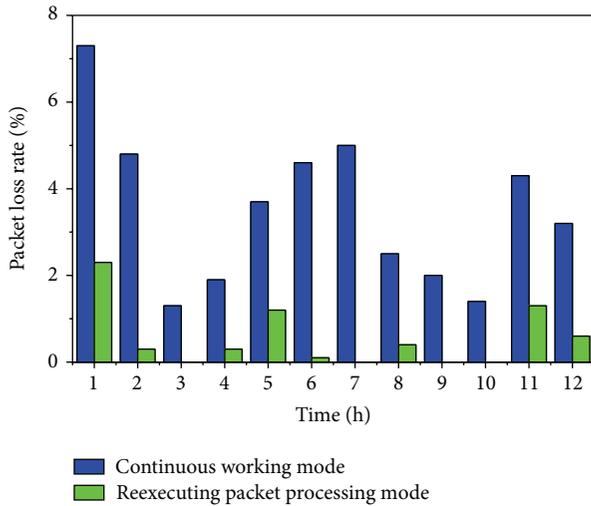


FIGURE 13: Packet loss analysis with different working mode.

2 seconds and the packet loss is reduced to 0.54%. Therefore, in the ZigBee wireless network, the scale data of sensor nodes are preprocessed by the gateway (or some controller nodes) which can effectively improve the transmission reliability of the system, optimize the wireless channel occupancy, and reduce the packet loss rate of the wireless network.

## 6. Concluding

In this paper, we proposed a ZigBee-based network scheme for home network applications. A stack structure node available for home service integration is provided. Based on the node, user can quickly develop a coverage monitoring application integration of ZigBee technology. User client server side adopts linear interpolation theory to evaluate the working status of each node in the ZigBee-based wireless network. A testbed of the proposed ZigBee-based wireless network system has been created to validate its basic functions. Experimental results of the demonstration project show that the testbed is convenient enough to perform wireless network coverage and monitoring tasks, and each node can visually display its working state.

## Competing Interests

The authors declared that they have no conflict of interests regarding this work.

## Acknowledgments

The research reported in this paper was carried out at the Remote Measuring and Control Lab, School of Instrument Science and Engineering, Southeast University, Nanjing, China. The authors would like to thank Kaijian Yin for help with the experiments. This work is partially supported by National Major Project (Grant nos. 2010ZX03006-006 and 2013ZX03001032-004), National 863 Program (Grant no. 2014AA01A702), National Natural Science Foundation

Project of China (Grant nos. 61272379 and 61325018), Jiangsu Provincial Key Technology R&D Program (Grant no. BE2012165), and the Ministry of Education Science and Technology Innovation Engineering Major Cultivation Project of China (Grant no. 107053).

## References

- [1] A. J. Jara, M. A. Zamora-Izquierdo, and A. F. Skarmeta, "Interconnection framework for mhealth and remote monitoring based on the internet of things," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 47–65, 2013.
- [2] S. F. Fang, L. D. Xu, Y. Q. Zhu et al., "An integrated system for regional environmental monitoring and management based on internet of things," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 2, pp. 1596–1605, 2014.
- [3] E. S. Nadimi, H. T. Sogaard, T. Bak, and F. W. Oudshoorn, "ZigBee-based wireless sensor networks for monitoring animal presence and pasture time in a strip of new grass," *Computers and Electronics in Agriculture*, vol. 61, no. 2, pp. 79–87, 2008.
- [4] W.-W. Chang, T.-J. Sung, H.-W. Huang et al., "A smart medication system using wireless sensor network technologies," *Sensors and Actuators A: Physical*, vol. 172, no. 1, pp. 315–321, 2011.
- [5] J. S. Byun, B. J. Jeon, J. Y. Noh, Y. Kim, and S. Park, "An intelligent self-adjusting sensor for smart home services based on ZigBee communications," *IEEE Transactions on Consumer Electronics*, vol. 58, no. 3, pp. 794–802, 2012.
- [6] Q. Hu, F. Li, and C.-F. Chen, "A smart home test bed for undergraduate education to bridge the curriculum gap from traditional power systems to modernized smart grids," *IEEE Transactions on Education*, vol. 58, no. 1, pp. 32–38, 2014.
- [7] S. Milo, S. Stefano, and N. Monica, "Wireless home automation networks for indoor surveillance: technologies and experiments internet of things for wireless and mobile communication," *EURASIP Journal on Wireless Communications and Networking*, vol. 2014, article 6, 2014.
- [8] S. K. Korkua and K. Thinsurat, "Design of ZigBee based WSN for smart demand responsive home energy management system," in *Proceedings of the 13th International Symposium on Communications and Information Technologies (ISCIT '13)*, pp. 549–554, IEEE, Surat Thani, Thailand, September 2013.
- [9] D. S. Kim, S.-Y. Son, and J. Lee, "Developments of the in-home display systems for residential energy monitoring," *IEEE Transactions on Consumer Electronics*, vol. 59, no. 3, pp. 492–498, 2013.
- [10] G. Tezel and Y. özbay, "A new approach for epileptic seizure detection using adaptive neural network," *Expert Systems with Applications*, vol. 36, no. 1, pp. 172–180, 2009.
- [11] C. Cervellera, D. Macciò, and M. Muselli, "Deterministic learning for maximum-likelihood estimation through neural networks," *IEEE Transactions on Neural Networks*, vol. 19, no. 8, pp. 1456–1467, 2008.
- [12] T. M. Jiang, X. Y. Li, J. Wan, and X. Liu, "A comprehensive assessment method for reliability enhancement testing based on D-S theory of evidence," in *Proceedings of the IEEE International Conference on Industrial Engineering and Engineering Management (IEEM '10)*, pp. 2375–2378, Macao, China, December 2010.
- [13] J. Zhao and S.-Q. Hu, "A new adaptive weighted fusion algorithm for multi-sensor tracking," in *Proceedings of the 1st International Conference on Machine Learning and Cybernetics*, pp. 285–287, Beijing, China, November 2002.

- [14] Y.-Z. Liu, J.-W. Zhang, and M.-B. Li, "A spatial-temporal fusion algorithm based support degree and self-adaptive weighted theory for multi-sensor," in *Proceedings of the International Conference on Machine Learning and Cybernetics (ICMLC '10)*, pp. 363–368, Qingdao, China, July 2010.
- [15] W.-T. Sung and J.-S. Lin, "Design and implementation of a smart LED lighting system using a self adaptive weighted data fusion algorithm," *Sensors*, vol. 13, no. 12, pp. 16915–16939, 2013.

## Research Article

# Length Variation Effect of the Impulse Response Model of a Secondary Path in Embedded Control

**Young-Sup Lee, Yunseon Choi, and Jeakwan Kim**

*Department of Embedded Systems Engineering, Incheon National University, 119 Academy-ro, Yeonsu-gu, Incheon 406-772, Republic of Korea*

Correspondence should be addressed to Young-Sup Lee; [yssl@inu.ac.kr](mailto:yssl@inu.ac.kr)

Received 25 December 2015; Accepted 25 February 2016

Academic Editor: Marco Anisetti

Copyright © 2016 Young-Sup Lee et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study presents theoretical and experimental investigation on the length variation effect of the impulse response function (IRF) for the secondary path model in active noise control using an embedded control board. A narrowband sweep noise was the disturbance for control in a duct with the length of 1800 mm. The IRF model incorporated into an adaptive feedforward filtered-x LMS (FxLMS) algorithm was then analyzed in the variation of its length in terms of the mean square error, computation complexity, stability requirement, and attenuation performance before and after control. The FxLMS algorithm with various IRF lengths was implemented in a dSPACE DS1104 embedded control board for the real-time control. Finally the most reasonable IRF length, considering the computation complexity and performance, can be determined through the systematic investigation. The results in this study can be used for practical active noise control systems.

## 1. Introduction

Active noise cancellation (ANC) is a technique based on adaptive feedforward control mostly to superpose an artificial sound to an unwanted disturbance noise [1–3]. The artificial sound, which is generated through the secondary loudspeaker, has to be of the same amplitude but the exact antiphase to the primary noise for best performance. An ANC system requires a good model for the compensation of the secondary path (or the plant) when the filtered-x least mean square (FxLMS) algorithm is used. It is known that the accuracy of the secondary path model has a decisive influence along with the convergence coefficient on the performance, stability, and convergence speed of the ANC system [1, 4]. The impulse response function (IRF) of the secondary path of the FxLMS approach is usually implemented as an FIR (finite impulse response) filter in a real-time embedded control board because the FIR filter is stable and easy to implement. An IIR (infinite impulse response) filter can be used to replace a long FIR filter in order to minimize the computation requirement, although the IIR filter can become rather easily unstable than an FIR one.

An optimum length of the IRF is critically important in terms of the computation complexity, performance, and stability of the control. The optimum length of the IRF must contain the main acoustic modes to be controlled or suppressed in the secondary path. The difference between the designed IRF and the actual secondary path causes the existence of the uncertainty of the plant [1, 5]. This difference is increased when the plant is perturbed. Thus the length of the IRF of the secondary path is critically important in terms of stability and performance in the FxLMS algorithm.

For robust stability of the FxLMS algorithm, Ren and Kumar suggested an SPR (strictly positive real) condition by combining the secondary path and its model [6]. Fraanje et al. investigated minimum norm regularization filters which can stabilize the update algorithm to prevent instability as the performance of the algorithm can be degraded not only by the plant model uncertainty but also by an unstable update method [7]. Berkhoff examined an approach for improved stability of adaptive feedforward controllers without measuring the plant uncertainty [8].

In this study, the length variation of the IRF of the secondary path is considered to analyze the control performance,

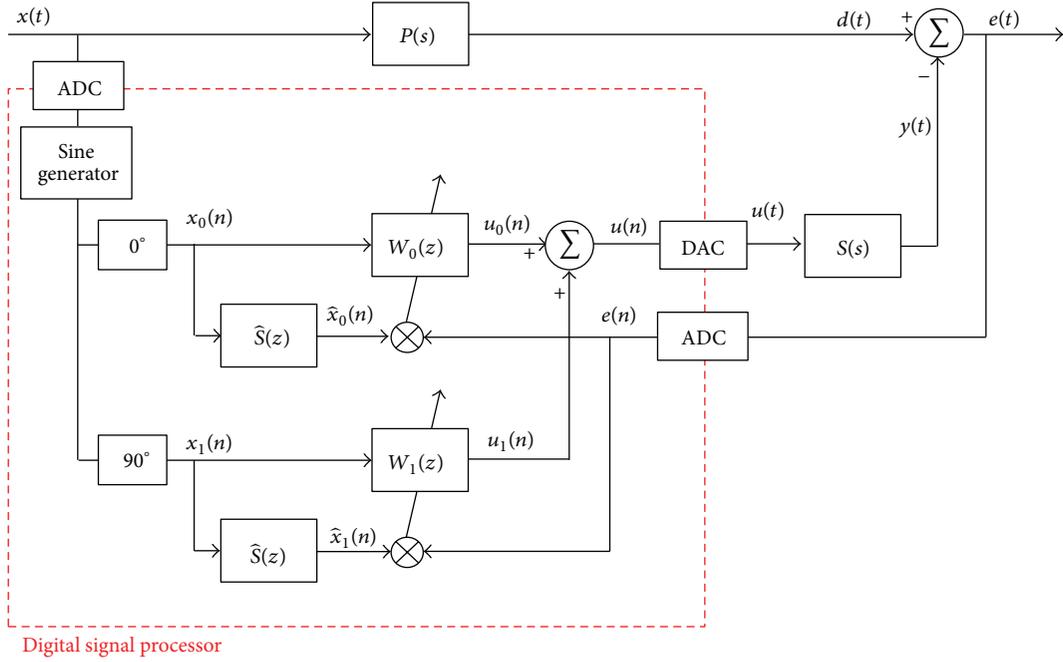


FIGURE 1: Block diagram of the narrowband FxLMS algorithm.

stability, and the computation complexity for a duct ANC system and an optimum IRF length is determined.

In Section 2, a theoretical examination is investigated for the effect of the IRF of the secondary path in the FxLMS algorithm. Section 3 shows the experimental setup for real-time ANC control of the duct system. Section 4 includes the results and analysis of the control experiment. Finally conclusions are described in Section 5.

## 2. Theoretical Considerations

Figure 1 shows a typical ANC block diagram of the SISO (single input single output) adaptive feedforward FxLMS algorithm that cancels the primary narrowband noise by the secondary source [1].

As illustrated in Figure 1,  $x(t)$  is the reference signal,  $d(t)$  is the disturbance signal,  $u(t)$  is the control effort,  $y(t)$  is the plant output and  $e(t)$  is the error signal, and  $P(s)$  and  $S(s)$  are the primary and secondary paths, respectively. The blocks inside the rectangle with red dashed line are implemented in a DSP (digital signal processor) where  $W(z)$  and  $\hat{S}(z)$  are the adaptive filter and the IRF of the secondary path model, respectively. In addition  $x(n)$ ,  $\hat{x}(n)$ ,  $u(n)$ , and  $e(n)$  are discrete signals to represent the reference, filtered reference, control effort, and error sequences, respectively.

Assuming 1 reference signal, 1 primary loudspeaker, 1 secondary loudspeaker, and 1 error microphone are installed in a duct system, the filtered reference sequence  $\hat{x}(n)$  can be expressed as

$$\hat{\mathbf{x}}_i(n) = \hat{\mathbf{S}}^T \mathbf{x}_i(n - \ell), \quad (i = 0, 1, \ell = 0 \sim L - 1), \quad (1)$$

where  $L$  is the IRF filter length of  $\hat{S}(z)$  and  $\hat{\mathbf{s}}$  is the coefficient of the IRF filter. The update equation of an adaptive feedforward narrowband FxLMS algorithm can be written by [1–4]

$$\mathbf{w}_i(n+1) = \mathbf{w}_i(n) + \alpha \hat{\mathbf{x}}_i(n) e(n), \quad (2)$$

where  $\alpha$  is the convergence coefficient. The computation to obtain the control effort  $u(n)$  in the DSP can be given by

$$u(n) = w_0(n) x_0(n) + w_1(n) x_1(n). \quad (3)$$

Provided that  $\hat{S}(z)$  is approximately identical to  $S(z)$ , the cancellation of the primary noise can be achieved during control as the IRF of the secondary path model can compensate for the actual secondary path. However, the IRF with some error makes a certain level of residual noise during control as

$$e(t) = d(t) - y(t) \neq 0. \quad (4)$$

Thus the accuracy of the IRF modeling of the secondary path is critically important in the implementation of an FxLMS based ANC system.

The computation complexity of the SISO FxLMS algorithm by length variation of the IRF filter can be given by Table 1, where the IRF filter length and control order number are  $L$  and  $C$ , respectively.

## 3. Experimental Setup

The experimental setup for the narrowband SISO ANC control is displayed in Figure 2. A duct (length = 1800 mm) ANC system is comprised of a primary loudspeaker at

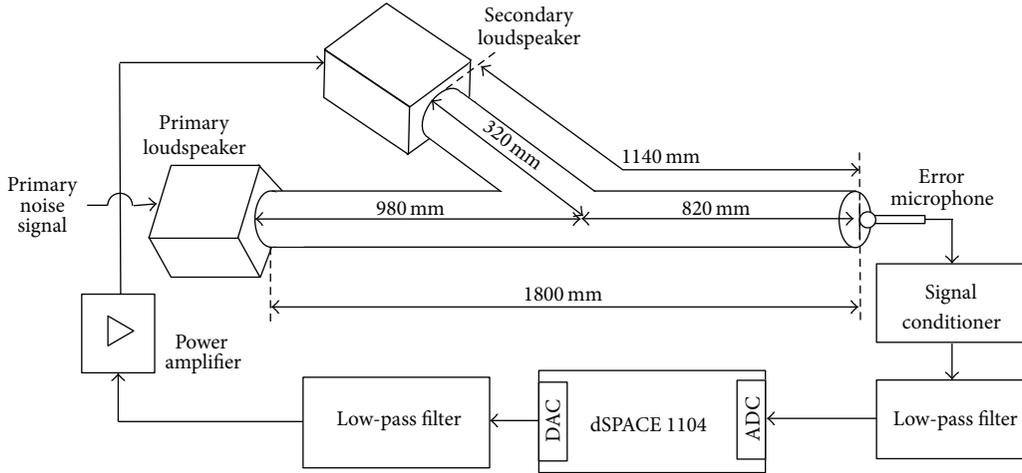


FIGURE 2: Experimental setup for the narrowband ANC for a duct.

TABLE 1: Computation complexity of the narrowband FxLMS algorithm.

Equation	Number of additions	Number of multiplications
(1)	$2(L-1)C$	$2LC$
(2)	$2C$	$2C+1$
(3)	$2C-1$	$2C$
Total	$2C(L+1)-1$	$2C(L+2)+1$

the left end, a secondary source (control loudspeaker), an error microphone (1/2" PCB 377B11) at the right end, a power amplifier (B&K 2716C), a PCB signal conditioner, two low-pass filters (cutoff frequency:  $f_c = 500$  Hz), and a dSPACE 1104 controller for the implementation of the SISO FxLMS algorithm.

The primary disturbance was a narrowband signal with some sinusoidal waves such as a powertrain interior noise of a passenger car. In this study, the largest four orders by the powertrain of the passenger car are the primary noise. The primary and secondary paths were measured at first and, after carrying out the offline modeling of the secondary path, the IRF model with various lengths was calculated and implemented in the control board. Then the real-time ANC control experiment was accomplished with each IRF model.

## 4. Results and Discussions

**4.1. IRF Modeling from the Measured Secondary Path.** White noise was used to measure the secondary path (secondary loudspeaker-error microphone) response when the sampling frequency was  $f_s = 6,000$  Hz. The sampling frequency was maintained the same all the time during this study. Figure 3(a) demonstrates that the IRF model from the measured secondary path response of the duct system has two peaks at 25 and 55 samples, respectively, and then decays quickly after 100 samples. The first and second peaks at the sample of 25 and 55, respectively, are the most important contributions to the IRF and indicate the time delays in

the secondary path of about 4.17 msec and 9.17 msec, respectively. As can be seen from Figure 3(a), at the sample of 400, the IRF settles down on zero and this allows assuming that the IRF length of 400 samples may represent the secondary path response accurately. Thus the digital plant model with the IRF length of  $L = 400$ ,  $\hat{S}_{400}(z)$ , can be an approximated model of the actual plant (secondary path),  $S(z)$ , as follows:

$$\hat{S}_{400}(z) \approx S(z). \quad (5)$$

Four different IRF lengths of  $L = 30, 63, 93,$  and  $400$  are considered and plotted in Figure 3(b) to compare their frequency response functions (FRFs). The magnitude of the FRF with  $L = 30$  (dashed lines) traces simply the averaged plot of the FRF with  $L = 400$  (thick solid lines) without resonances or antiresonances in the frequency range below 1200 Hz and its phase response looks almost as a straight line. For  $L = 63$ , the magnitude response follows some main resonances and antiresonances as shown in Figure 3(b). In the case of  $L = 93$ , the magnitude and phase responses are very similar to those of  $L = 400$ .

The mean square error (MSE) is defined by the difference between the two plant models of  $\hat{S}_{400}(z)$  and  $\hat{S}_L(z)$  when the input signal  $x(n)$  is a white noise and can be represented as

$$\text{MSE} = E \left[ x(n) \left[ \hat{S}_{400}(z) - \hat{S}_L(z) \right] \right], \quad (6)$$

where  $E[\ ]$  is the expectation operator. As the MSE indicates the accuracy of the secondary path model by comparing the outputs of the actual secondary path and its model against a white noise input, the MSE is an important factor to assess the model reliability.

The MSE curve against the IRF length of  $L = 0-120$  is plotted in Figure 4(a) and it indicates that the MSE is dramatically reduced by the increase of the IRF length. Actually the inclusion of the peaks in the IRF of Figure 4(a) gives the most critical influences to the accuracy of modeling of the secondary path. For the best performance and robust stability, every peak is necessary in the IRF model; however

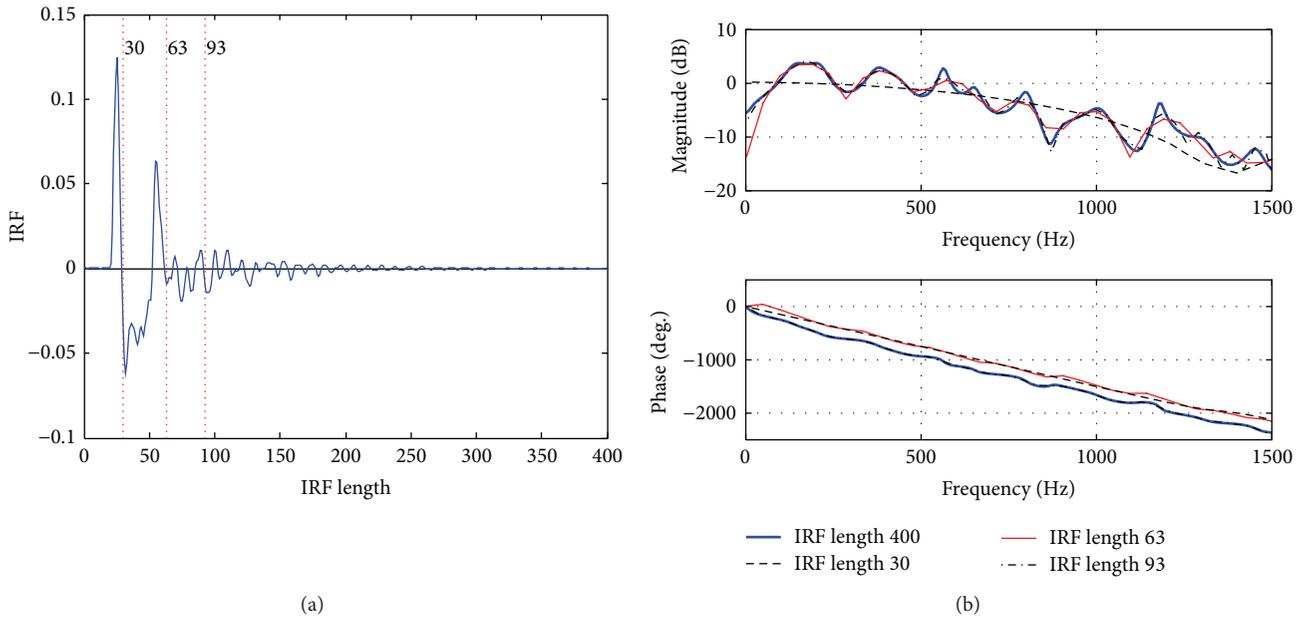


FIGURE 3: Plant modeling from the measured secondary path response. (a) IRF. (b) Comparison of FRFs for four different IRF lengths at  $L = 30, 63, 93,$  and  $400$  (top: magnitude, bottom: phase).

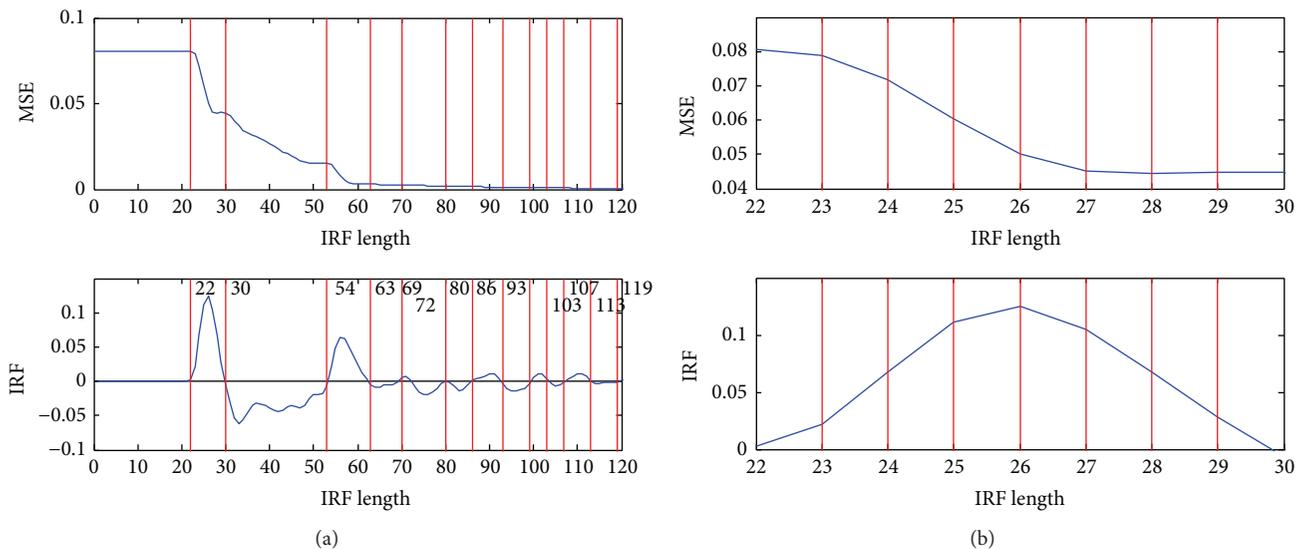


FIGURE 4: MSE against variation of the IRF length. (a) IRF length of  $L = 0-120$ . (b) IRF length of  $L = 22-30$ .

this can increase the computation complexity. Thus the IRF length must be chosen carefully.

There are 13 vertical lines in Figure 4(a) which implies 13 different IRF sample numbers where the IRF graph passes 0. The MSE below  $L = 22$  where there is no peak in the IRF maintains its maximum level. As the IRF length increases after  $L = 22$ , the MSE decreases accordingly as shown in Figure 4(a). Especially at  $L > 60$  the MSE is very low compared to that of  $L = 22$ .

In addition, the change of the MSE against a peak in the IRF graph at  $L = 22-30$  is investigated in detail as illustrated in Figure 4(b) and it demonstrates that the MSE reaches

the lowest level as soon as  $L$  is the next sample at  $L = 27$  after the peakiest sample at  $L = 26$ . This is important information in determining the minimized IRF length for the FxLMS algorithm particularly when the allowance of the computation complexity is limited.

**4.2. Comparison of Control Performance by IRF Length Variation.** Figure 5 displays the comparison of the spectrograms before and after controls with the FxLMS algorithm in the duct system of Figure 2 when the IRF length is 63 samples against a narrowband disturbance which is defined as a 3-second sweep sine signal with four different orders of CI

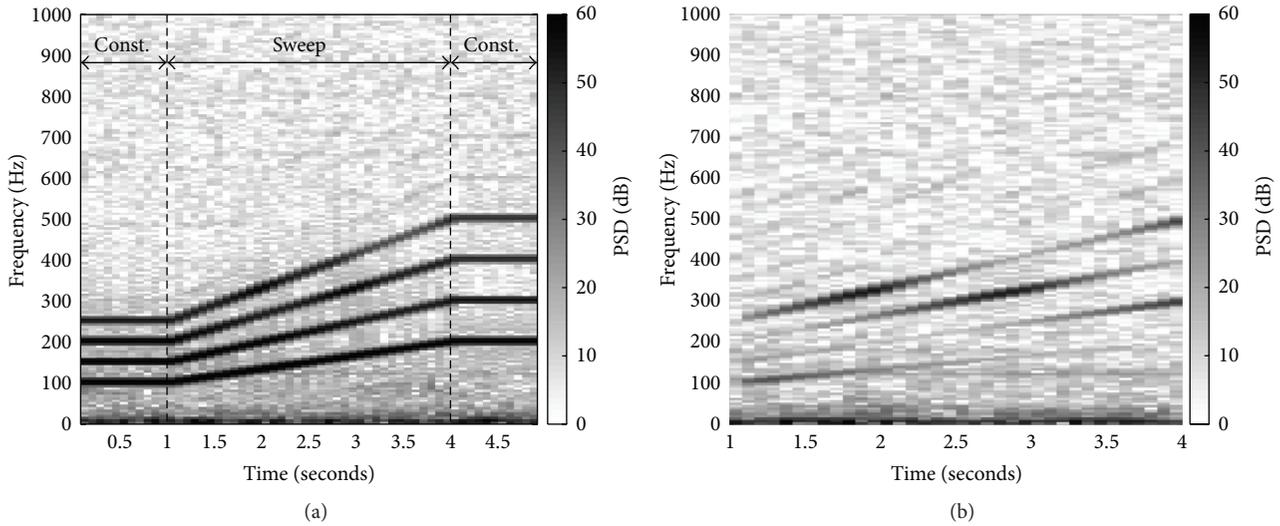


FIGURE 5: Comparison of the spectrograms before and after controls against a narrowband disturbance. (a) Spectrogram before control; (b) spectrogram after control (IRF length  $L = 63$ ).

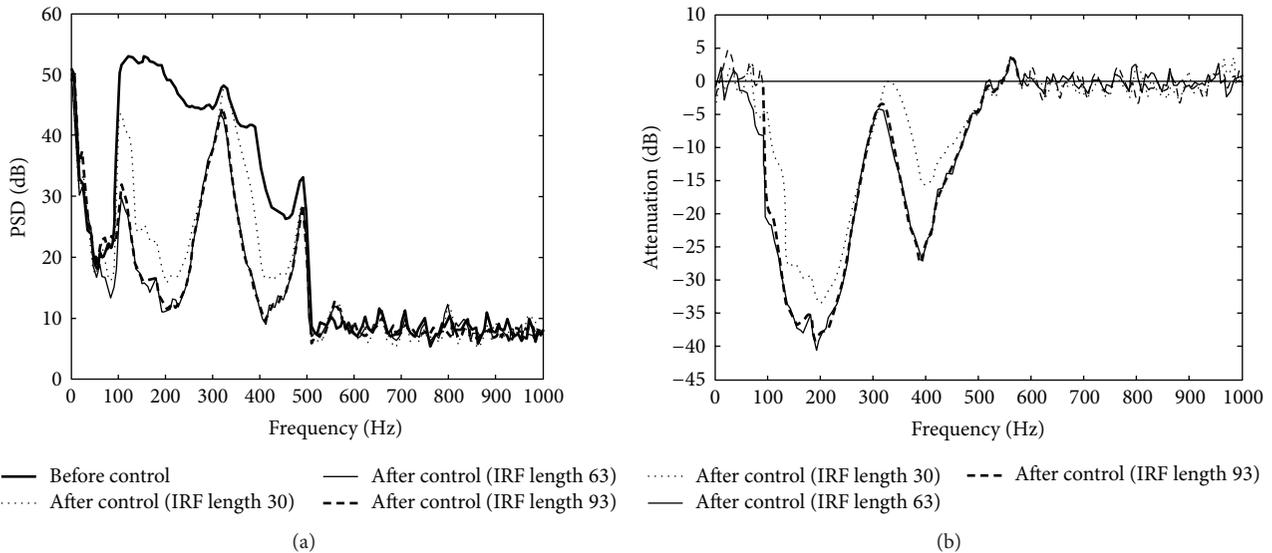


FIGURE 6: Comparison of narrowband sweep real-time control performances in terms of error spectra and attenuation between before control (thick lines) and after control (IRF length  $L = 30$ : dotted lines,  $L = 63$ : thin lines, and  $L = 93$ : thick dashed lines). (a) Error PSD. (b) Attenuation.

(100–200 Hz), C1.5 (150–300 Hz), C2 (200–400 Hz), and C2.5 (250–500 Hz).

The convergence coefficient  $\alpha$  defined in (2) was 0.01 for the control. The spectrogram plotted in Figure 5(a) shows that the duration of the disturbance is 5 seconds consisting of 1 second at the beginning for constant signals, 3 seconds for sweep signals, and another 1 second at the end for constant signals. As can be seen from Figure 1, the disturbance signal  $d(t)$  was measured by the error microphone when the primary loudspeaker generated the narrowband signal  $x(t)$  in the duct before control.

Figure 5(b) shows the spectrogram for the sweep duration (1–4 seconds) only after control with the FxLMS algorithm, as defined in Figure 1, when the IRF length of  $L = 63$

and indicates that an eminent reduction for every order is achieved. There are the *dark* parts and *light* parts of each line (order) after control in Figure 5(b). The dark and light parts indicate less attenuation and more attenuation, respectively. The more attenuated frequency ranges are 100–250 Hz and 350–450 Hz as plotted in Figures 5(b) and 6(b). Any order in the two frequency ranges was attenuated and shown lightly.

The reduction levels (dB) in the power spectral density (PSD) after control are plotted in Figure 6 by the IRF length variation. Figure 6(a) presents the PSD comparison before control (thick solid lines) and after control for the three different IRF lengths of  $L = 30$  (dotted lines), 63 (thin solid lines), and 93 (thick dashed lines) samples. The PSDs after control reveal that the longer the IRF offers the further

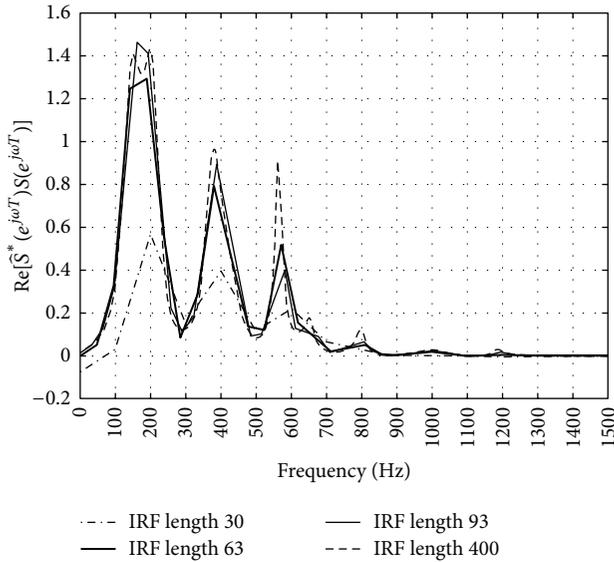


FIGURE 7: Robustness of the secondary path by the variation of the IRF length.

the reduction; however the reduction difference between the IRF lengths of  $L = 63$  and  $93$  samples is not clearly distinguished. The largest reduction achieved about 25–40 dB at 100–250 Hz when both  $L = 63$  and  $93$ , but the reduction at  $L = 30$  was 5–10 dB smaller than them as illustrated in Figure 6(b).

#### 4.3. Comparison of Robust Stability by IRF Length Variation.

It is noted from Figure 6(b) that the attenuations at some frequency ranges, such as about 100–120, 300–350, and 480–500 Hz, are not attained properly compared to other frequencies. This suggests that, in these frequency ranges, the control stability could be easily collapsed if the convergence coefficient is increased by even a small value to achieve further attenuation. This is basically caused by the acoustical properties of the duct and can be additionally explained by the primary and secondary paths.

It is known that the stability requirement for the FxLMS algorithm can be expressed by [1, 5, 6]

$$\operatorname{Re} \left[ S(e^{j\omega T}) \hat{S}_L(e^{j\omega T}) \right] > 0, \quad \forall \omega T, \quad (7)$$

where  $\omega$  is the frequency and  $T$  is the sampling interval. The stability requirement over frequency in (7) is satisfied with three different IRF lengths of  $L = 63$  (thick solid line),  $93$  (thin solid line), or  $400$  (dashed line) as presented in Figure 7.

But, in the case of  $L = 30$  (dashed and dotted line), the requirement is unsatisfied at the frequencies below 100 Hz and this implies the FxLMS control with  $L = 30$  can cause instability if the control frequency range involves below 100 Hz. Thus the stability requirement in (7) must be considered in the determination of the IRF length in advance of control.

Figure 7 displays three outstanding peaks, assuming that  $\operatorname{Re}[S(e^{j\omega T})\hat{S}_L(e^{j\omega T})] > 0.2$ , at about 100–270, 320–460, and 530–600 Hz are the three largest resonant frequencies of the

FRF of the secondary path in Figure 3(b) and two valleys between them at about 270–320 and 460–530 Hz. In case of  $L = 30$ , the frequency ranges for the peaks and valleys are partially different. The frequencies of the valleys such as around 300 Hz and 500 Hz in Figure 7 are almost identically located to the unattenuated frequencies as can be seen from Figure 6.

This coincidence discloses maintaining as higher values of  $\operatorname{Re}[S(e^{j\omega T})\hat{S}_L(e^{j\omega T})]$  as possible in the frequency range of interest and this allows higher convergence coefficients to achieve further attenuation. The valleys near zero in Figure 7 strictly restrict the increase of the convergence coefficients and this limits the control performance. This is because the increased convergence coefficient can cause more performance but less stability and if it is excessive the control system can be unstable. Thus obtaining robust stability with higher performance needs higher values of  $\operatorname{Re}[S(e^{j\omega T})\hat{S}_L(e^{j\omega T})]$  at the valleys in the frequency ranges of interest.

#### 4.4. Comparison of Computation Complexity by IRF Length Variation.

Computation complexity defined in Table 1 against various IRF lengths with the FxLMS algorithm their mean values of attenuation in the duct ANC experiment are summarized in Table 2. The IRF lengths considered in the control experiment are  $L = 30, 54, 63, 70, 80, 86, 93, 99$ , and  $400$ . Table 2 shows the attenuation is almost the same of about  $-21.50$  dB in average when  $L$  is greater than or equal to 63. Considering the computation complexity and performance by the IRF length variation, it indicates that the IRF length of  $L = 63$ , which contains the two major peaks in the IRF graph as plotted in Figure 3(a), is the most reasonable choice in the duct system. This analysis allows relieving the computation complexity by reducing about 84% of the IRF length with the same performance as presented in Table 2.

Therefore, the IRF length variation effect of the secondary path model  $\hat{S}_L(z)$  is investigated in this study in terms of MSE, involvement of major peaks in the IRF graph, stability requirement, and attenuation and these terms have strong correlations with the optimum IRF length. The results in this study can be useful in various ANC applications.

In case the primary noise is broadband, it is expected that this approach can be worked well since it is based on the analysis of the secondary path.

## 5. Conclusions

This study presents the length variation effect of the secondary path IRF model in an ANC duct system. The main investigations can be described in the following.

The major peaks in the full IRF graph are needed to be involved in the IRF model for the FxLMS algorithm. Although the computation complexity is limited, the IRF model requires containing the peakiest and the next samples at least of the last major peak in the full IRF graph. The stability requirement of  $\operatorname{Re}[S(e^{j\omega T})\hat{S}_L(e^{j\omega T})] > 0$  must be considered in the determination of the IRF length in

TABLE 2: Computational complexity per sample and mean values of attenuation by the IRF length variation ( $\alpha = 0.01$ ).

IRF length ( $L$ )	Number of additions	Number of multiplications	Mean value of attenuation [100~500 Hz]
30	247 (7.70%)	257 (7.99%)	-15.01 dB
54	439 (13.69%)	449 (13.96%)	-20.78 dB
<b>63</b>	<b>511 (15.93%)</b>	<b>521 (16.20%)</b>	<b>-21.48 dB</b>
70	567 (17.68%)	577 (17.94%)	-21.57 dB
80	647 (20.17%)	657 (20.42%)	-21.37 dB
86	695 (21.67%)	705 (21.91%)	-21.63 dB
93	751 (23.42%)	761 (23.66%)	-21.21 dB
99	799 (24.91%)	809 (25.15%)	-21.37 dB
400	3207 (100%)	3217 (100%)	—

advance of control. In addition, robust stability with higher performance demands higher values of  $\text{Re}[S(e^{j\omega T})\hat{S}_L(e^{j\omega T})]$  at the valleys in the frequency ranges of interest.

This investigation clarifies that the most reasonable IRF length, considering the computation complexity and performance, for the implementation in an embedded control board can be determined systematically through the investigation on an ANC duct system in this study. For future study, this approach will be extended to an ANC system against a broadband primary noise.

## Competing Interests

The authors declare that they have no competing interests.

## Acknowledgments

This work was supported by the Incheon National University Research Grant in 2012.

## References

- [1] S. J. Elliott, *Signal Processing for Active Control*, Academic Press, New York, NY, USA, 2001.
- [2] K. Lee, W. S. Gan, and S. M. Kuo, *Subband Adaptive Filtering: Theory and Implementation*, John Wiley & Sons, 2009.
- [3] Y. Kajikawa, W.-S. Gan, and S. M. Kuo, "Recent advances on active noise control: open issues and innovative applications," *APSIPA Transactions on Signal and Information Processing*, vol. 1, article e3, 2012.
- [4] S. Ryu, Y. J. Park, and Y.-S. Lee, "Active suppression of narrow-band noise by multiple secondary sources," *Journal of Sensors*, vol. 2016, Article ID 6276828, 9 pages, 2016.
- [5] M. Morari and E. Zafiriou, *Robust Process Control*, Prentice-Hall, London, UK, 1989.
- [6] W. Ren and P. R. Kumar, "Adaptive active noise control: structures, algorithms and convergence analysis," in *Proceedings of the Engineering for Environmental Noise Control (Inter-Noise '89)*, pp. 435–440, Newport Beach, Calif, USA, November 1989.
- [7] P. R. Fraanje, S. J. Elliott, and M. Verhaegen, "Robustness of the filtered-X LMS algorithm—part II: robustness enhancement by minimal regularization for norm bounded uncertainty," *IEEE*

*Transactions on Signal Processing*, vol. 55, no. 8, pp. 4038–4047, 2007.

- [8] A. P. Berkhoff, "A technique for improved stability of adaptive feedforward controllers without detailed uncertainty measurements," *Smart Materials and Structures*, vol. 21, no. 6, Article ID 064003, 2012.

## Research Article

# A Smart Gateway Architecture for Improving Efficiency of Home Network Applications

Fei Ding,<sup>1,2</sup> Aiguo Song,<sup>3</sup> En Tong,<sup>1,2</sup> and Jianqing Li<sup>3</sup>

<sup>1</sup>*School of Information Science and Engineering, Southeast University, Nanjing 210096, China*

<sup>2</sup>*R&D Center, China Mobile Group Jiangsu Co., Ltd., Nanjing 210029, China*

<sup>3</sup>*School of Instrument Science and Engineering, Southeast University, 2 Sipailou, Nanjing, Jiangsu 210096, China*

Correspondence should be addressed to Aiguo Song; [a.g.song@seu.edu.cn](mailto:a.g.song@seu.edu.cn)

Received 26 September 2015; Accepted 29 November 2015

Academic Editor: Gwanggil Jeon

Copyright © 2016 Fei Ding et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A smart home gateway plays an important role in the Internet of Things (IoT) system that takes responsibility for the connection between the network layer and the ubiquitous sensor network (USN) layer. Even though the home network application is developing rapidly, researches on the home gateway based open development architecture are less. This makes it difficult to extend the home network to support new applications, share service, and interoperate with other home network systems. An integrated access gateway (IAGW) is proposed in this paper which upward connects with the operator machine-to-machine platform (M2M P/F). In this home network scheme, the gateway provides standard interfaces for supporting various applications in home environments, ranging from on-site configuration to node and service access. In addition, communication management ability is also provided by M2M P/F. A testbed of a simple home network application system that includes the IAGW prototype is created to test its user interaction capabilities. Experimental results show that the proposed gateway provides significant flexibility for users to configure and deploy a home automation network; it can be applied to other monitoring areas and simultaneously supports a multi-ubiquitous sensor network.

## 1. Introduction

The IoT (Internet of Things) is considered one of the major communication advances in recent years, since it offers the basis for the development of independent cooperative services and applications [1]. Extensive research is underway to explore the potential applications using this concept in different areas, such as a ZigBee-based, wearable scheme for monitoring physiological parameters [2], monitoring animal presence and pasture time over an extended area [3], the management of medication and health care [4], and, in particular, home network applications (HNAs). The IoT potential for HNAs is reported in [5–8]. The implementation of IoT for monitoring the environmental conditions of a home is proposed in [5]. A smart home testbed based on the pedagogical model of project-based learning (PBL) for undergraduate education is proposed in [6]. A wireless home automation network for indoor surveillance is presented in [7]. A synchronous smart and wireless home energy management

scenario was proposed in [8] and [9], respectively. Each of these studies identifies potential capacities for self-configuration, comprehensive management, and communication capabilities.

For HNAs, the home gateway is the most important sink node (SN) that provides access and management capabilities for the child nodes (CNs), such as home devices or smart home terminals. Many articles about ZigBee-based home gateways have been published recently [10–12]. Authors in [10] propose a ZigBee-based homecare gateway that deals with the collection of terminal medical data and eventually passes the data to the doctor for remote diagnosis. Authors in [11] introduce a household controller supporting both ZigBee and infrared communication. An OSGi-based architecture for the dynamic integration of ZigBee home networks into home gateways is proposed in [12], where ZigBee home network devices are represented as device proxy service bundles. However, according to these HNA scenarios, data interaction between SN and CN is achieved mostly by proprietary

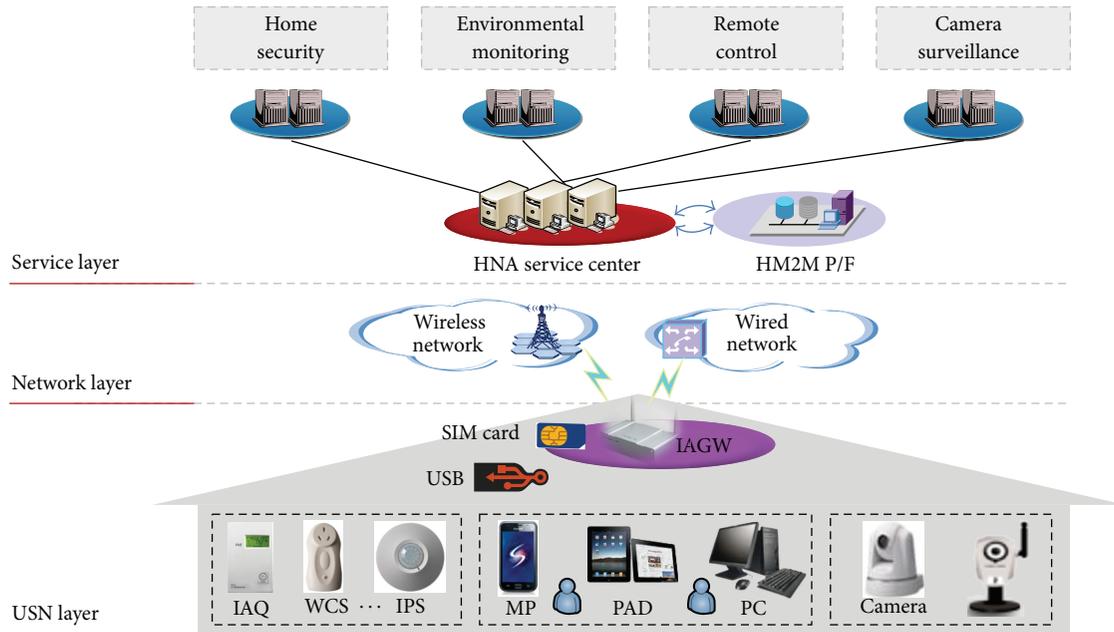


FIGURE 1: A M2M platform based home network architecture.

wireless protocol, and both SN and CN are specially designed for an HNA scheme, and, as such, lack interoperability. These proprietary gateways and their systems are difficult to extend so as to support new services, integrate new data, or interoperate with other HNAs.

Machine-to-machine platform (M2M P/F) integrated service schemes which offer self-configuration and expandability have significant potential for improving the efficiency of traditional home automation systems [13]. In [14], a snapshot is proposed of the latest progress in M2M standardization of things such as requirements, architecture, candidate protocols, security aspects, and device management. In [15], an IoT-based retractable bollard management system is described to govern vehicular access appropriated for restricted urban areas. In [16], biomedical applications and various low power wireless sensor and link technologies are discussed, but the scheme integration proposed is only suited to the BlackBerry system. In [17], M2M architecture for applications in healthcare, energy management, and entertainment is proposed, and the performance tradeoffs inherent in existing designs are evaluated.

In this paper, an integrated access gateway (IAGW) architecture is proposed with standard interfaces for supporting various application nodes in home environments, ranging from on-site configuration to node and service access. Both the upward and downward communication interfaces of the IAGW are standardized. The IAGW based home network system architecture adopts a separation principle between terminal management and service implementation; the former is realized by the M2M P/F, and the latter is provided by the home service layer. The working mode and parameter configuration of the IAGW gateway can be set using a web console. If the configuration of the IAGW is completed correctly, it will allow management by the M2M

P/F automatically. Meanwhile, the system service data are directly transferred to the home service layer through the IAGW. Based on the proposed IAGW and the M2M P/F fusion architecture, the home sensors and services can be standardized and easily managed. Thus, it can reduce the redevelopment workload and promote the interoperability between the multi-HNA systems. The system architecture, the design methods of the IAGW based HNA, and the demonstration experiment will be discussed in Sections 2–4.

## 2. System Overview

As shown in Figure 1, the hierarchical architecture of the home network system consists of 3 layers: the ubiquitous sensor networks (USN) layer, the network layer and the service layer.

The main role of the IAGW is data transmission between the CNs and the network layer. The CNs in the USN layer include indoor air quality (IAQ), wireless control socket (WCS), infrared proximity sensor (IPS), video camera, and other sensors or actuators in the HNA. Except for the video camera, the other CNs are integrated with ZigBee modules and network ability. The video camera connects to the IAGW gateway through a Wi-Fi wireless interface or an RJ45 wired interface. Users enjoy the convenience and safety of HNA services through the personal computer (PC), mobile phone (MP), and PAD. The network layer consists of a wired network and a 3G/4G wireless cellular network. The service layer includes an HNA service center (HNASC) and a home M2M platform (HM2M P/F); the former is responsible for home services management, while the latter provides interfaces to the IAGW and supports gateway management. The HNASC supports receiving, authenticating, and storing real-time data collected by the IAGW. It also connects with the

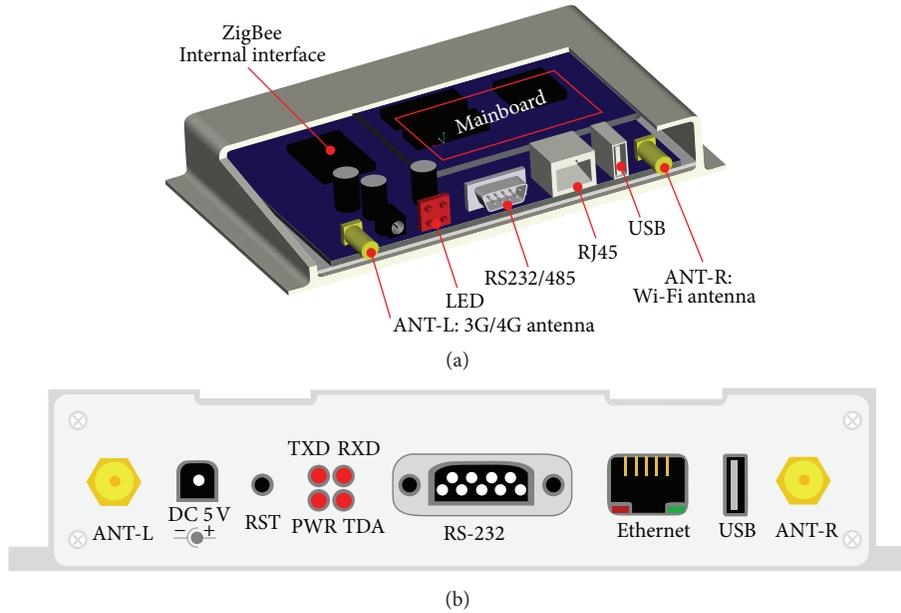


FIGURE 2: Hardware architecture of the proposed gateway. (a) Stack structure of the proposed gateway. (b) Front view of the gateway interface side.

business operating support system (BOSS) to realize operations like business order synchronization and push information to users in the form of hand-held terminal and short message service (SMS).

### 3. Design Methods

**3.1. IAGW Prototype.** Figure 2 shows the stack structure hardware architecture of the proposed IAGW. The IAGW prototype is designed to provide standard access to the CNs deployed in the HNA system. The mainboard is the core unit of the IAGW that is designed with a MCU, a SDRAM, and a NAND flash; it runs Linux 2.6 or a later operating system (OS). The USB interface is used to connect a memory card for local data storage, especially video monitoring data. The 4 LED indicators of the IAGW represent different working states, corresponding to power on, network connect/disconnect, data transmission, and alarm on/off, respectively.

The IAGW proposes an onboard 22-pin internal interface and an RS232/485 external interface connected to the mainboard, which are convenient for ZigBee SN integration. If user needs to connect and control other wireless standard home devices, the corresponding SN can be redeveloped and integrated through the two interfaces. The dimensions of the onboard internal interface that adjusts to the ZigBee SN are shown in Figure 3.

**3.2. Functional Architecture.** The functional architecture of the IAGW is shown in Figure 4. The USN layer interface is realized through the ZigBee radio and its networking ability. The public network interface mandates an IP-based network, such as 2G/3G/4G, or Ethernet. The USN Manager

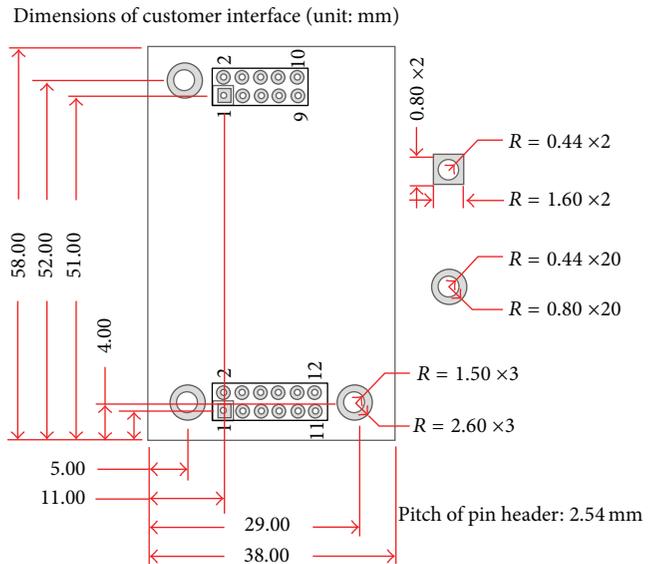


FIGURE 3: Dimensions of the onboard internal interface (unit: mm).

module manages the home USN, which consists of providing a uniform network interface to the packet manager or other modules, detection of USN network status (connection status OK/NG), and acquisition of all of the USN network service types through the USN I/F.

The “Registration” module is used to manage node registration to the IAGW and request gateway registration to the HM2M P/F. Meanwhile, the “Gateway Configuration” module provides the communication parameters and working mode configuration of the IAGW through HM2M P/F.

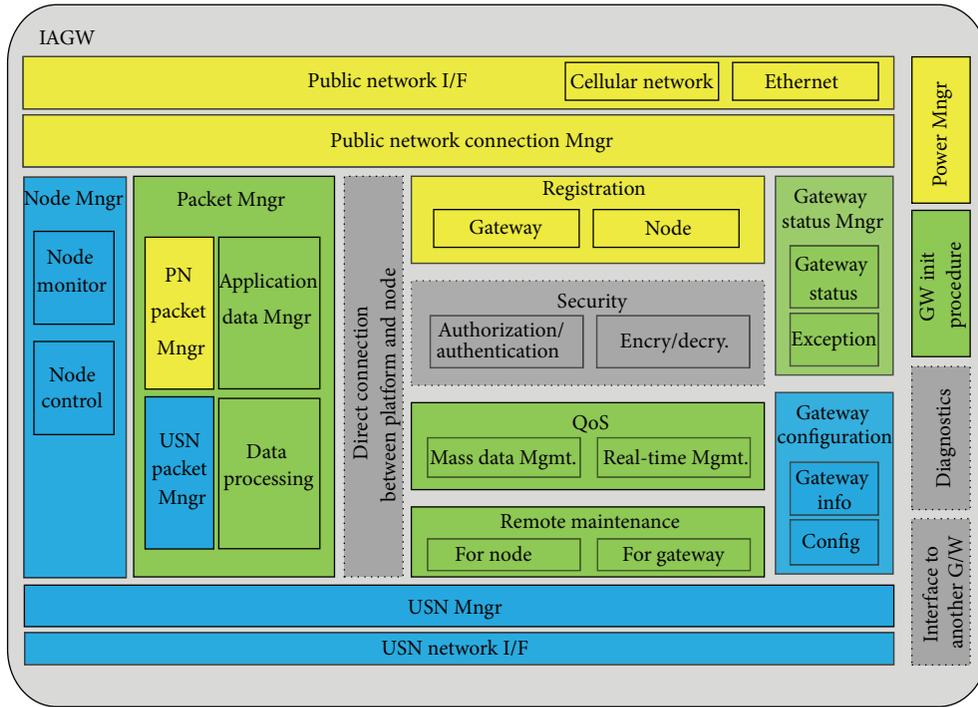


FIGURE 4: Functional architecture of the proposed IAGW gateway.

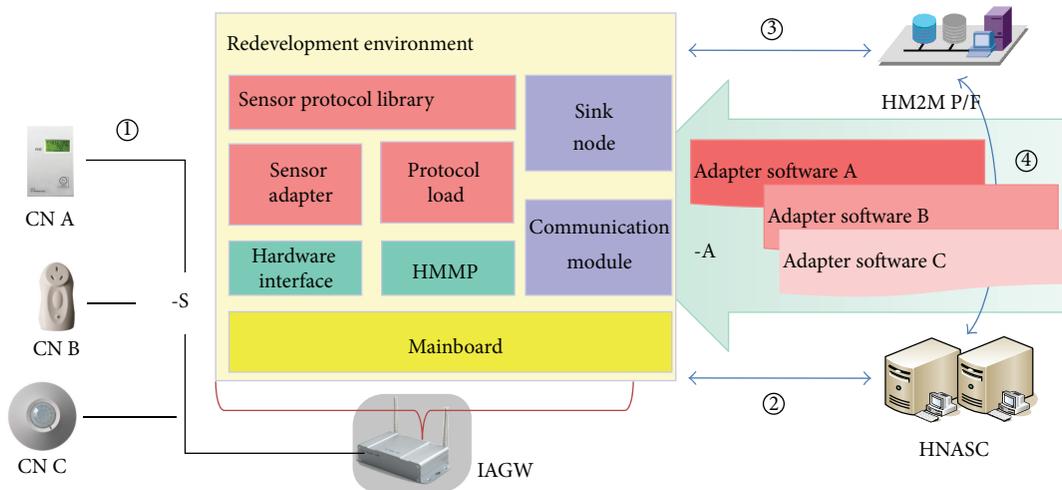


FIGURE 5: Interaction architecture of the IAGW based home network application.

The “Node Mngr” module manages CNs, including aspects such as node ID management, node status monitoring, and node control. The node monitor is a submodule of the node manager that monitors the node’s status. The “Packet Mngr” module of the proposed IAGW repackages data received from CNs to a standard form and sends them to HNASC P/F, parses requests initiated by the platform, and packages them into the appropriate form to send to the CNs. The “Security” module includes authorization/authentication and en/decryption, and the “QoS (Quality of Service)” module

guarantees service quality. The “Remote Maintenance” module supports remote maintenance, such as IAGW and SN software upgrades from the HM2M P/F.

3.3. Adaptive Access and M2M Business Interaction Method. As shown in Figure 5, the IAGW based HNA system can be divided into 2 kinds of communication interfaces: the HMMP- (Home M2M protocol-) S and HMMP-A.

The HMMP-S interface (step ①) realizes a ZigBee-based wireless communication standard protocol between the CNs

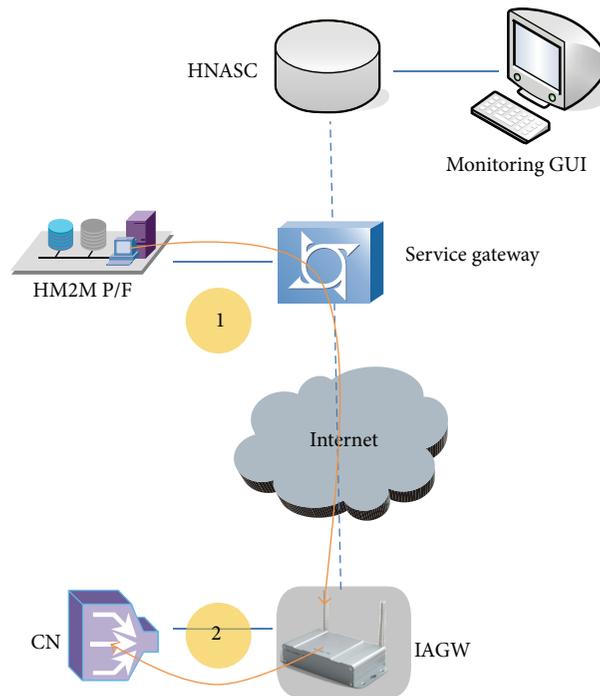


FIGURE 6: Configuration procedure of the home terminal.

and the IAGW, so as to solve the access of the various types of home terminals. It relates to home network attributes, reading the CNs' ID and description information and setting the communication port of the ZigBee SN. It relates to home network services, acquiring all of the home network service types through the IAGW. The HMMP-A protocol (steps ②, ③, and ④) belongs to the data transmission interface between the IAGW and the HM2M P/F. The HMMP-S interface protocol is forwarded directly to the HMMP-A interface via the IAGW, and HMMP-A protocol includes the packet of the HMMP-S. Meanwhile, if the IAGW connects to the HNASC platform, the HMMP-A protocol should be the expansion of the HMMP-S protocol, so that the reliability of communication between the IAGW and the HM2M P/F, HM2M P/F, and HNASC can be secured.

The HM2M P/F uses *CONFIG\_SET* command to configure the various home terminals. As shown in Figure 6, the adapter software according to a home terminal transmits through the HM2M P/F (step 1) and remote push to the IAGW (step 2).

Aiming at CN node management based on HMMP-S protocol, the IAGW takes appropriate action based on the status, such as adjusting the data sending rate or the error status. Meanwhile, it reports node status to the HM2M P/F through the IAGW, allowing to manage the status regularly. HMMP-S protocol also implements control of the CNs' working behaviors, and node control commands are initially issued by the IAGW or HNASC P/F. Home node data acquisition and processing flow are shown in Figure 7: step 0 represents the configuration of the home CN, including data acquisition and data reporting rules; step 1 represents the home node using sensor adapter protocol, which is loaded by

the HM2M P/F to read data; step 2 indicates that the home node uses *TRANSPARENT\_DATA* command to transmit data to the HNASC P/F; step 3 presents the hand-held monitoring and displays CN sensing data and status.

**3.4. User Development Procedure.** Figure 8 shows the user development step and interface in the proposed HM2M P/F based home network system. By configuring the communication port of the IAGW, the communication between the IAGW and SN is secured. After port configuration is finished, it would be autosaved to the internal memory of the IAGW.

If the HMMP-S communication configuration is finished, then the communication parameters of the HMMP-A can be set. The main parameters that should be configured include the IP address, the port, and the secret keys of the HNASC P/F. The IAGW also can be configured in encryption mode. When the *COMM\_CONFIG* data are executed, the IAGW sends a communication key request to the HM2M P/F, which is automatically assigned by the latter. Between the CNs and the IAGW, also the IAGW and the M2M P/F, monitoring and maintenance of communication are implemented with *HEART\_BEAT* and *HEART\_BEAT\_ACK* packets. Hence, the data transmission link of the HNA system has been established; the HM2M P/F can implement remote management for the IAGW and CNs, and through the HNASC P/F to support kinds of home services for user.

## 4. System Implementation

**4.1. Prototype Gateway Based HNA.** As shown in Figure 9, a testbed of a home network system for validating the basic functions of the proposed smart IAGW has been created. The



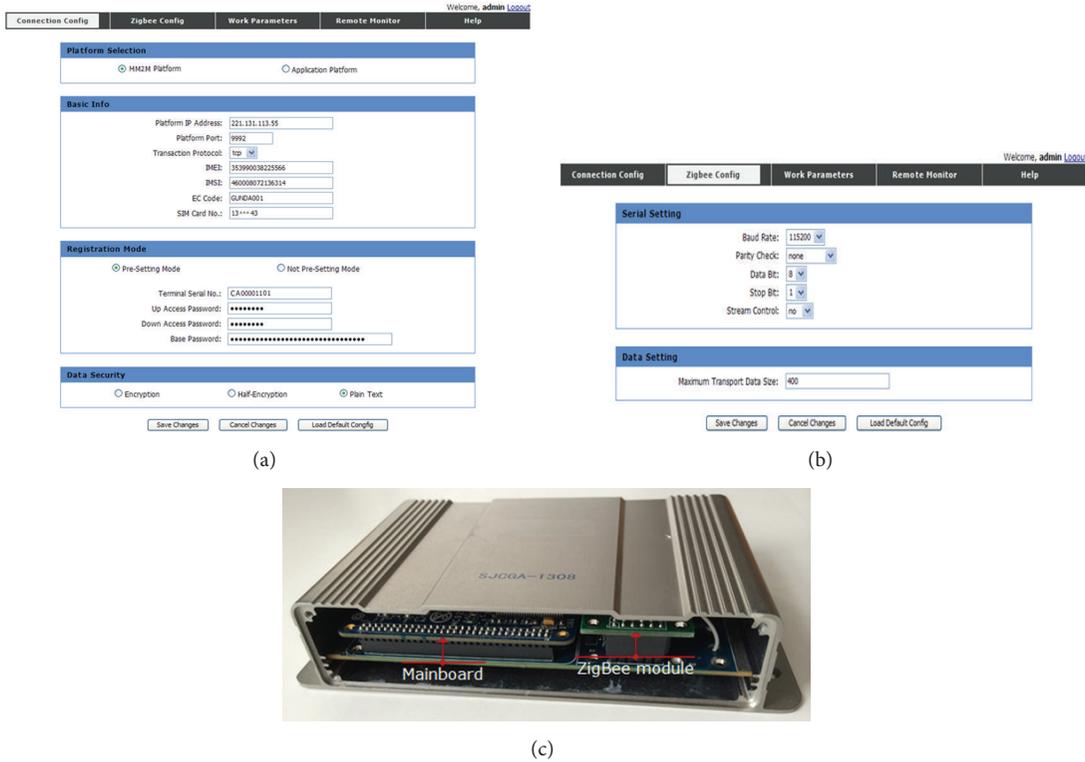


FIGURE 10: Parameters setting GUIs of the prototype IAGW with the web console. (a) Connection configuration GUI. (b) ZigBee communication configuration GUI. (c) Side view of the proposed prototype gateway.



FIGURE 11: Digital map navigation GUI for monitoring platform.

IAGW and the ZigBee SN are implemented according to the ZigBee configuration options.

Figure 11 shows the local monitoring client of the home network system. The GUI functions include data acquisition and management, graphic displays, person-machine dialogue, and real-time and historical database storage. User can control every device in the home network conveniently by using this service client. Three standard monitoring scenes for the home network system, including away monitoring, disarming mode, and a stop alarm, are set for choosing. User

can conveniently select the working mode of the system. Once an alarm occurs, it will automatically occur in the PC client and locate the alarm event in the event map GUI. The hand-held remote client can receive the alarm SMS notification; it also has remote access to the home camera, making it convenient for the user to confirm the alarm.

User can remotely check the home environment parameters in real-time and access live video through the hand-held MPs. The GUI running on the user MPs consists of 5 modules, including warning information, node management, real-time data, service instruction, and system configuration. As shown in Figure 12, (a) is the indoor temperature monitoring chart and (b) is the monitoring video of the proposed testbed. The former uses different colors to distinguish the temperature alarm interval, which makes it convenient for user to understand whether it exceeds a set threshold. The latter uses the Wi-Fi network attached to the user's MPs to acquire real-time monitoring video and ensure home safety.

4.2. *Demonstration Experiment.* Some tests are carried out to evaluate the stability of our IAGW based HNA system. The IAQ node in the proposed testbed senses the home environment in real time; the temperature, humidity, and CO<sub>2</sub> data are collected and sent to the HNASC P/F through the IAGW at intervals of 5 seconds. The tests last for 24 hours, and the data are stored and visualized at the HNASC P/F in real time. The sample results of the home environmental monitoring experiments are shown in Figure 13. At 17:10 and 4:10, the large curve fluctuations are caused by blowing air to



FIGURE 12: Graphic user interfaces running on hand-held terminals. (a) Indoor temperature GUI. (b) Remote video access GUI.

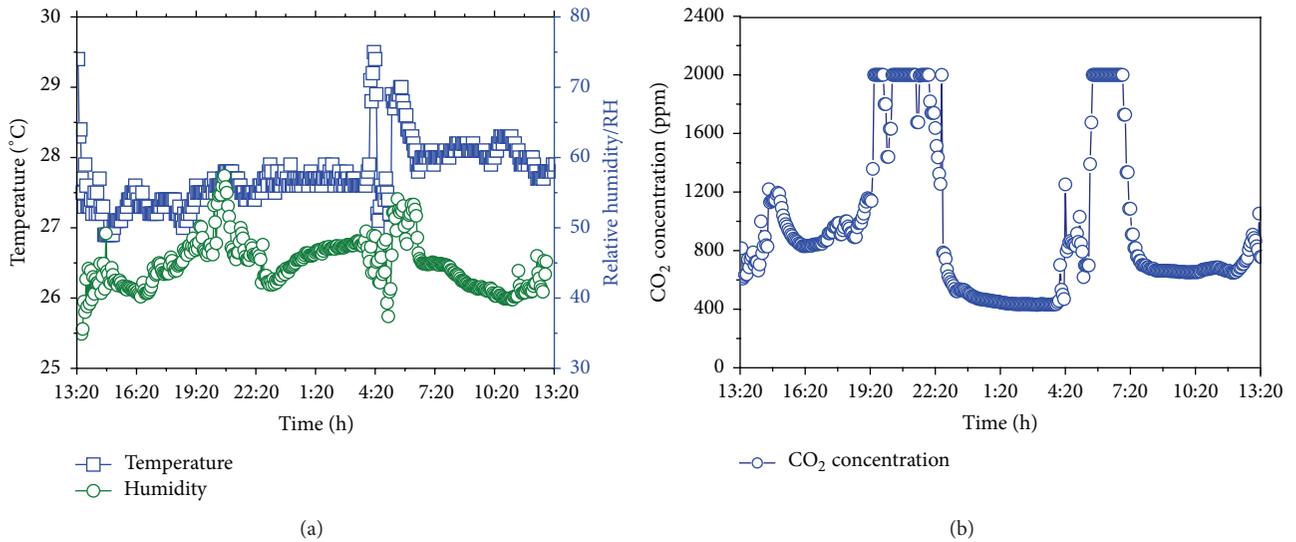


FIGURE 13: Sample results of the indoor environmental monitoring experiments in 24 hours: (a) temperature and humidity and (b) CO<sub>2</sub>.

TABLE 1: The comparison between the traditional gateway and the IAGW.

ID	Characteristic	Conventional gateways [2, 3, 8–10]	IAGW
1	Multi-interface	Mainly support ZigBee radio and have poor expansion	Using the stack hardware structure, it supports ZigBee, Wi-Fi, and Bluetooth radio and is easy to expand
2	Secondary development ability	Need a large amount of work	New development workload is relatively small
3	Communication link monitoring	Need to add custom development	Automatic support after system development
4	System openness	Based on terminal development, all interfaces are private	The interfaces are standard and the openness and source utilization of the system are enhanced

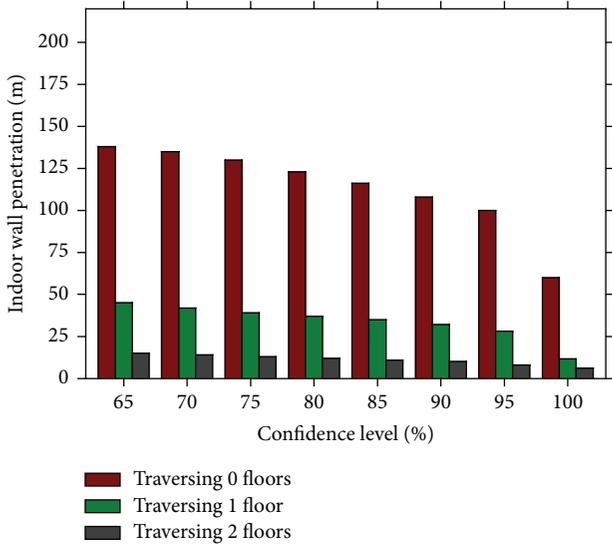


FIGURE 14: Node-to-node spacing capability.

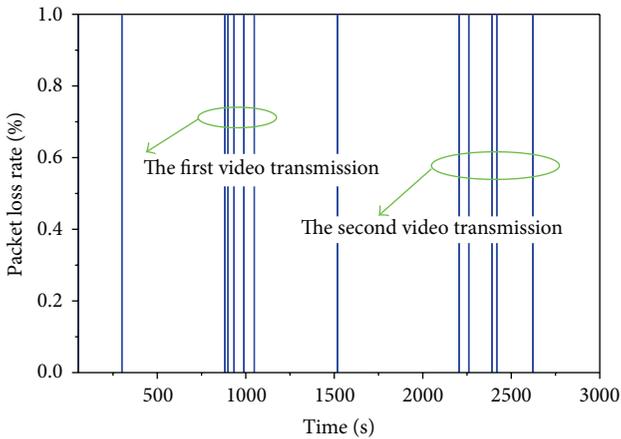


FIGURE 15: Performance evaluation between the two same frequency technologies.

an IAQ node, and the room temperature, humidity, and CO<sub>2</sub> changes during one day are displayed in these figures.

Figure 14 plots node-to-node spacing versus confidence level for situations where there is no directional antenna and no line-of-sight present. The testing method is carried out by continuously broadcasting the control data and receiving the successfully returned signal via the ZigBee network. It is certain that when a link is set up inside a building, it will exhibit a shorter range than that suggested by the free-space guidelines, even when a line-of-sight is present. In cluttered environments, the range reduction can be very significant. In fact, radio waves will propagate through brick walls, concrete floors, and plasterboard partitions, among other materials, but a loss will be incurred (as compared with free-space propagation).

The IAGW supports both ZigBee and Wi-Fi 2.4 GHz band wireless standards. In our proposed HNA system, the demand for ZigBee-based data transmission is much larger

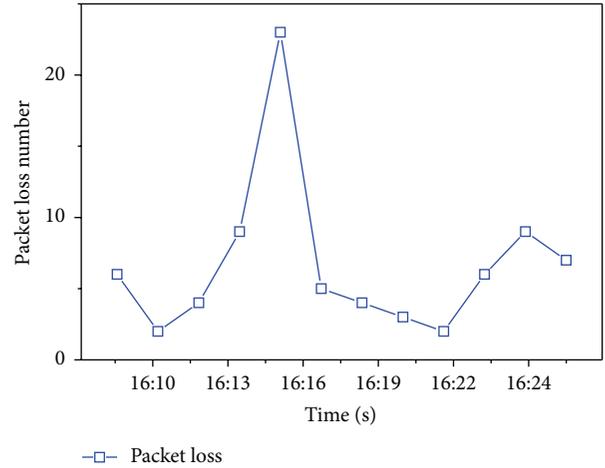


FIGURE 16: The packet loss monitoring based on the HM2M P/F.

than that for Wi-Fi communication. The influence of Wi-Fi-based video transmission on ZigBee wireless communication is shown in Figure 15. When the home network system only uses ZigBee wireless data transmission, the system packet loss rate is about 0.4%; when the user starts the remote video monitoring task, the system packet loss rate reaches about 7.6%. Hence, in the HNA system, the user needs to balance video access with system data packet loss. The optimal working mode of the HNA system is to minimize the number of video access services unless an alarm occurs. This can improve system reliability effectively.

Based on the HM2M P/F, the remote data interaction between the IAGW and the USN layer can be monitored in real time. Figure 16 indicates the periodically monitoring results of packet loss in a live cellular network, which helps to evaluate the wireless communication quality and emergency response. In poor wireless network condition, the failed transmitted test data of the USN layer will be cached in the IAGW integrated storage and will be submitted to the HNASC in the next data transmission process with the normal network link working stage.

### 5. Conclusions

A smart gateway architecture for improving the efficiency of home network applications is proposed in this paper. The gateway has stack architecture and provides on-board standard internal and external interfaces which make wireless communication module integration convenient. Among the gateway and various home sensor nodes, as well as between the gateway and the service layer, the gateway realizes the interface standard. Therefore, it can easily realize the standardization of the home network system. The hardware and business management logic separation of the gateway improve the scalability of the system. The working mode and parameter configuration of the gateway can be set through a web browser. After the completion of the configuration, the system will directly enter the work mode properly, which can reduce the development workload of new home network applications. Experimental results show that the

proposed gateway successfully realizes home network applications using little infrastructure, and it provides a faster and more flexible approach for building and deploying home automation networks.

### Conflict of Interests

The authors declared that they have no conflict of interests regarding this work.

### Acknowledgments

This work is partially supported by National Major Project (Grant no. 2010ZX03006-006), National 863 Program (Grant no. 2014AA01A702), National Natural Science Foundation Project of China (Grant nos. 61272379 and 61325018), Jiangsu Provincial Key Technology R&D Program (Grant no. BE2012165), and the Ministry of Education Science and Technology Innovation Engineering Major Cultivation Project of China (Grant no. 107053). The authors would like to thank Ke-fa Wang and Kai-jian Yin for help with the experiments.

### References

- [1] A. J. Jara, M. A. Zamora-Izquierdo, and A. F. Skarmeta, "Interconnection framework for mHealth and remote monitoring based on the internet of things," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 47–65, 2013.
- [2] K. Malhi, S. C. Mukhopadhyay, J. Schnepfer, M. Haefke, and H. Ewald, "A zigbee-based wearable physiological parameters monitoring system," *IEEE Sensors Journal*, vol. 12, no. 3, pp. 423–430, 2012.
- [3] E. S. Nadimi, H. T. Sogaard, T. Bak, and F. W. Oudshoorn, "ZigBee-based wireless sensor networks for monitoring animal presence and pasture time in a strip of new grass," *Computers and Electronics in Agriculture*, vol. 61, no. 2, pp. 79–87, 2008.
- [4] W.-W. Chang, T.-J. Sung, H.-W. Huang et al., "A smart medication system using wireless sensor network technologies," *Sensors and Actuators A: Physical*, vol. 172, no. 1, pp. 315–321, 2011.
- [5] S. D. T. Kelly, N. K. Suryadevara, and S. C. Mukhopadhyay, "Towards the implementation of IoT for environmental condition monitoring in homes," *IEEE Sensors Journal*, vol. 13, no. 10, pp. 3846–3853, 2013.
- [6] Q. Hu, F. Li, and C.-F. Chen, "A smart home test bed for undergraduate education to bridge the curriculum gap from traditional power systems to modernized smart grids," *IEEE Transactions on Education*, vol. 58, no. 1, pp. 32–38, 2015.
- [7] M. Spadacini, S. Savazzi, and M. Nicoli, "Wireless home automation networks for indoor surveillance: technologies and experiments Internet of Things for wireless and mobile communication," *EURASIP Journal on Wireless Communications and Networking*, vol. 2014, no. 1, p. 6, 2014.
- [8] S. K. Korkua and K. Thinsurt, "Design of ZigBee based WSN for smart demand responsive home energy management system," in *Proceedings of the 13th International Symposium on Communications and Information Technologies (ISCIT '13)*, pp. 549–554, September 2013.
- [9] D. S. Kim, S.-Y. Son, and J. Lee, "Developments of the in-home display systems for residential energy monitoring," *IEEE Transactions on Consumer Electronics*, vol. 59, no. 3, pp. 492–498, 2013.
- [10] H. Y. Tung, K. F. Tsang, H. C. Tung, K. T. Chui, and H. R. Chi, "The design of dual radio ZigBee homecare gateway for remote patient monitoring," *IEEE Transactions on Consumer Electronics*, vol. 59, no. 4, pp. 756–764, 2013.
- [11] I.-K. Hwang, D.-S. Lee, and J.-W. Baek, "Home network configuring scheme for all electric appliances using ZigBee-based integrated remote controller," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 3, pp. 1300–1307, 2009.
- [12] Y.-G. Ha, "Dynamic integration of zigbee home networks into home gateways using OSGI service registry," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 2, pp. 470–476, 2009.
- [13] G. Lopez, P. Moura, J. I. Moreno, and A. De Almeida, "ENERSip: M2M-based platform to enable energy efficiency within energy-positive neighbourhoods," in *Proceedings of the IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPs '11)*, pp. 217–222, IEEE, Shanghai, China, April 2011.
- [14] J. Swetina, G. Lu, P. Jacobs, F. Ennesser, and J. S. Song, "Toward a standardized common M2M service layer platform: introduction to oneM2M," *IEEE Wireless Communications*, vol. 21, no. 3, pp. 20–26, 2014.
- [15] L. Foschini, T. Taleb, A. Corradi, and D. Bottazzi, "M2M-based metropolitan platform for IMS-enabled road traffic management in IoT," *IEEE Communications Magazine*, vol. 49, no. 11, pp. 50–57, 2011.
- [16] S. Adibi, "Link technologies and BlackBerry mobile health (mHealth) solutions: a review," *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 4, pp. 586–597, 2012.
- [17] M. Chen, J. F. Wan, S. Gonzalez, X. F. Liao, and V. C. M. Leung, "A survey of recent developments in home M2M networks," *IEEE Communications Surveys and Tutorials*, vol. 16, no. 1, pp. 98–114, 2014.

## Research Article

# Energy Harvesting and Information Transmission Protocol in Sensors Networks

Xue-Fen Zhang<sup>1</sup> and Chang-Chuan Yin<sup>2</sup>

<sup>1</sup>College of Information Technology, Beijing Union University, Beijing 100101, China

<sup>2</sup>Key Laboratory of Universal Wireless Communications, Ministry of Education, Beijing University of Posts and Telecommunications, Beijing 100876, China

Correspondence should be addressed to Xue-Fen Zhang; [zhangxuefen@buu.edu.cn](mailto:zhangxuefen@buu.edu.cn)

Received 17 June 2015; Revised 2 September 2015; Accepted 7 September 2015

Academic Editor: Gwanggil Jeon

Copyright © 2016 X.-F. Zhang and C.-C. Yin. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We focus on the design of transmission protocol for energy harvesting wireless sensors. The sensors can harvest the energy from the environment, but they cannot charge and discharge at the same time. We propose a protocol for energy harvesting and wireless transmission, which contains two steps. In the first step, the sensor harvests the energy from environment, and the energy harvesting rate is controlled by the harvested energy power of the energy saving device (ESD). In the second step, some data should be transmitted to the receiver in a certain time. Considering one slot time, the first part of the time is devoted exclusively to energy harvesting, and the remaining time of the slot is for transmitting the information data. Assume that  $Q$  bits are transmitted to the receiver within one time slot; we establish the relationship between the harvesting energy time and the transmit data time. In addition, we analyze the system outage probability performance over the Rayleigh fading channel.

## 1. Introduction

The energy harvesting in wireless sensor network has become increasingly attractive with sustainable and long lasting power supplies. With development of IOT, more and more sensors are embedded in medical device or building structures to sense the environment [1–3]. Sensor nodes are powered by batteries that often cannot be replaced because of the inaccessibility of the devices (e.g., inside the human body). Once the battery of a sensor node is exhausted, the node dies. In this case, a solution may be to harvest the energy from the environment. The energy harvesting methods include absorbing energy from the solar wind, vibration, RF radio waves, temperature differences, and airflow, thermoelectric effects, and other physical phenomena [4].

Among all these methods, radio signal radiated by ambient transmitters became a new source for wireless energy harvesting, and harvesting energy from ambient RF signals can power a wireless sensors network [5]. Besides, recent advance in designing highly efficient rectifying antennas will enable more efficient wireless energy harvesting from RF

signals in the near future [6, 7]. More recently, the study of wireless powered communication networks (WPCNs) has received a lot of interest. For a WPCN, the energy harvested from ambient RF signals is used to power wireless terminals in the network [8]. In addition, the advantage of RF solution lies in the fact that RF signals can carry energy and information at the same time, and such an approach can also reduce the cost of communication networks, since peripheral equipment to take advantage of external energy sources can be avoided. Besides, ambient RF from TV and cellular communications is widely available in urban areas (day and night, indoors and outdoors) [9].

There is some progress in the recent research of energy harvesting. For example, the information theoretic capacity with an energy harvesting transmitter was discussed in [10]. The paper introduced two different capacity-achieving schemes and concluded that the capacity of an energy harvesting AWGN channel is the same as that of a traditional AWGN channel as long as the average available energy is the same. In [11], the problem of transmission time minimization in an energy harvesting setting for a point-to-point

communication system was studied. In [12, 13], the authors investigated the minimization of the transmission time for a given amount of data through power control based on known energy arrivals over all the time. In [14], the authors considered the power allocation for an access-controlled transmitter with energy harvesting capability based on causal observations of the channel fading state. In addition, the problem of transmission time minimization was also solved by mapping it to the problem of throughput maximization.

In this work, we consider the design of transmission protocol for energy harvesting wireless sensor system, in which the users transmit their independent information using their individually harvested energy in the uplink. Our work is unlike traditional research on Simultaneous Wireless Information and Power Transfer (SWIPT), which assumes the simultaneous energy and information transmissions to sensors (or users) in the downlink. A major problem of this assumption is that practical circuits cannot charge and discharge at the same time. We assume that the transmitter needs to deliver  $Q$  bits within  $T$  seconds, where  $T$  is the duration of a transmission frame. The transmitter can only harvest energy or deliver information at any given time, but not both. As a result, the transmitter needs to decide when to switch between the transmission data (TD) mode and the energy harvesting (EH) mode. We derive the optimal mode switching rule at the transmitter to achieve various tradeoffs between the minimum transmission outage probability in TD mode and the average harvested energy in EH mode.

The rest of this paper is organized as follows. Section 2 provides the system model. We analyze the outage probability in Section 3 and establish the optimization problem in Section 4. Section 5 presents numerical examples. Finally, Section 6 concludes the paper.

## 2. System Model

In this paper, we assume the system includes one pair of single-antenna transmitter (Tx) and receiver (Rx). Suppose  $Q$  bits are transmitted to the receiver within  $T$  seconds, where  $T$  is the duration of a transmitter frame. In every frame, there are two phases of transmission. The first phase is called harvesting time, in which ESD collects the energy from the environment. Since the ESD cannot charge and discharge simultaneously, the second step is for the data transmission in the remaining frame time. Next, we give the key parameters in the communication protocol for energy harvesting and information processing at Tx. Define the save-ratio  $\rho$  as the fraction of the one transmission slot time, and, during the time interval  $[0, \rho T]$ , Tx harvest energy by ESD. The remaining slot time, from time  $\rho T$  to  $(1 - \rho)T$  seconds, is used for the information transmission. In this paper, the relation between  $\rho$  and outage probability is established.

Let  $h_{s,d}$  be the source-destination channel coefficients. The channel from the source to the destination is modeled as frequency nonselective fading and additive noise [1, 5, 7]. Therefore, the channel coefficients are assumed to be constant during a complete slot time and may vary from a slot time to another.

The receiver signal at the destination node in one slot time,  $y_r(t)$ , can be written as

$$y_r(t) = \sqrt{\frac{P_s}{d_{s,d}^\eta}} h_{s,d} x(t) + n_{s,d}, \quad (1)$$

where  $h_{s,d}$  is the source-to-receiver channel gain. We model  $h_{s,d}$  as independent zero-mean, circularly symmetric complex Gaussian random variables with variances  $\sigma_{s,d}^2$ , the envelope of which follows Rayleigh distribution (for Rayleigh fading the channel amplitude squares  $|h_{s,d}|^2$  are exponentially distributed). The noise is modeled as independent zero-mean, circularly symmetric complex Gaussian random variables with zero means and variances.  $d_{s,d}$  is the source-to-receiver distance,  $P_s$  is the average power transmitted from the source in the time slot,  $\eta$  is the path loss exponent, and  $x(t)$  is the source message with unit power; that is,  $E[|x(t)|^2] = 1$ , where  $E[\cdot]$  is the expectation operator and  $|\cdot|$  is the absolute value operator.  $n_{s,d}$  denotes Gaussian noise with unit variance.

According to the protocol, the total buffered energy at  $t = \rho T$ , which is the time of start to transmit data within a transmission slot, is given by  $E_T = P_T \rho T$ , where  $P_T$  is the harvested energy power of the ESD.

Denote  $P_s = E_T / (1 - \rho)T = P_T(\rho / (1 - \rho))$  as the average total power, which is constant over the entire transmission period.

## 3. Outage Probability

**3.1. SNR Analysis.** In this section, we analyze the instantaneous signal-to-noise ratio (SNR) and derive the probability density function (PDF) of the SNR.

The instantaneous SNR in the destination is given by

$$\gamma_0 = \frac{P_s |h_{s,d}|^2}{d_{s,d}^\eta N_0}, \quad (2)$$

where  $N_0$  is the variance of AWGN in the corresponding channels. We can see that the random variable  $\gamma_0 = P_s |h_{s,d}|^2 / d_{s,d}^\eta N_0$  is an exponential random variable with parameter  $\beta_0 = N_0 d_{s,d}^\eta / P_s \sigma_{s,d}^2$ . Hence, the PDF of  $\gamma_0$  is

$$f_{\gamma_0}(x) = \begin{cases} \beta_0 e^{-\beta_0 x}, & x > 0, \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

**3.2. Outage Probability Performance Analysis.** For a given  $h$ , the cluster-to-cluster amplify-and-forward relay network's mutual information is

$$I = \log_2(1 + \gamma_0). \quad (4)$$

The outage probability at Rx can be expressed as

$$\begin{aligned} P_{\text{out}} &= \Pr\{I < R\} = \Pr\{\log_2(1 + \gamma_0) < R\} \\ &= \Pr\{\gamma_0 < 2^R - 1\}, \end{aligned} \quad (5)$$

where  $R$  is the required target rate.

Because

$$f_{\gamma_0}(x) = \begin{cases} \beta_0 e^{-\beta_0 x}, & x > 0, \\ 0, & \text{otherwise,} \end{cases} \quad (6)$$

hence

$$\begin{aligned} P_{\text{out}} &= \Pr\{\gamma_0 < 2^R - 1\} = \int_0^{2^R - 1} f_{\gamma_0}(x) dx \\ &= \int_0^{2^R - 1} \beta_0 e^{-\beta_0 x} dx = 1 - e^{-\beta(2^R - 1)}. \end{aligned} \quad (7)$$

#### 4. Optimization Problem

The outage probability is an important criterion of the system performance. In this section, we determine an optimal save-ratio to improve the performance of outage probability. Suppose  $Q$  bits of data are generated and transmitted within a time slot of duration  $T$  seconds, and, for a given save-ratio  $\rho$ , the data delivery only takes place in the last  $(1 - \rho)T$  seconds of each time slot, which results in an effective rate of  $R \geq R_{\text{eff}} = Q/(1 - \rho)T$  bits/sec. In this section, we study the relation between  $\rho$  and outage probability.

Consider the following optimization problem:

$$\begin{aligned} \min P_{\text{out}} \\ \text{where } \rho \in (0, 1). \end{aligned} \quad (8)$$

According to (7), the optimization problem can be written as

$$\min_{\rho \in (0, 1)} \left(1 - e^{-\beta(2^{R_{\text{eff}}} - 1)}\right). \quad (9)$$

According to  $R \geq R_{\text{eff}} = Q/(1 - \rho)T$ ,  $\beta_0 = N_0 d_{s,d}^\eta / P_s \sigma_{s,d}^2$ , and  $P_s = E_T / (1 - \rho)T = P_T / (\rho / (1 - \rho))$ , we can rewrite the problem as

$$\begin{aligned} \min_{\rho \in (0, 1)} \left( 1 \right. \\ \left. - \exp\left(-\frac{N_0 d_{s,d}^\eta}{P_T (\rho / (1 - \rho)) \sigma_{s,d}^2} (2^{Q/(1-\rho)T} - 1)\right) \right). \end{aligned} \quad (10)$$

Hence, the optimum  $\rho$  is given by

$$\min_{\rho \in (0, 1)} \frac{(1 - \rho) N_0 d_{s,d}^\eta}{P_T \rho \sigma_{s,d}^2} (2^{Q/(1-\rho)T} - 1). \quad (11)$$

For (11), we can find the numerical solution based on golden section search and parabolic interpolation techniques [15].

#### 5. Numerical Examples

In this section, we evaluate the performance of the proposed transmission schemes. First, we will calculate the mutual information  $I$  and the outage probability versus the save-ratio

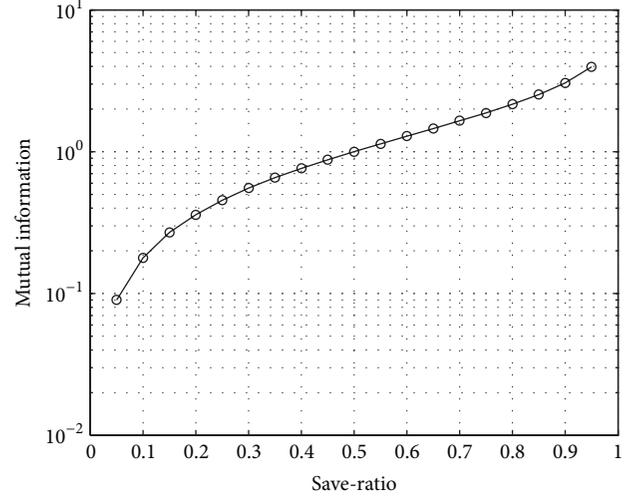


FIGURE 1: The system mutual information.

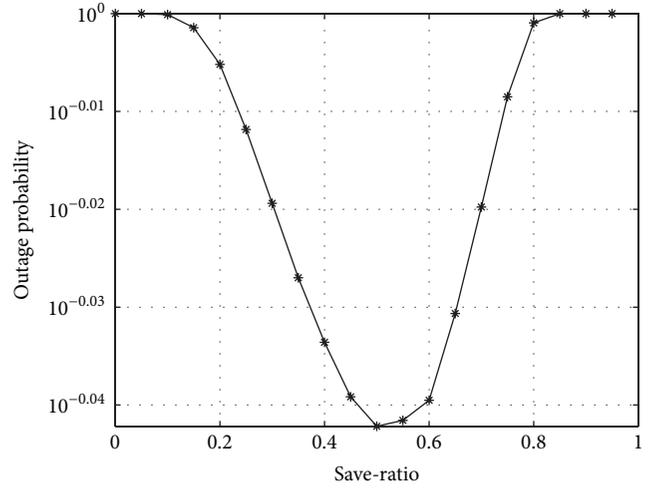


FIGURE 2: The relationship of outage probability and the save-ratio  $\rho$ .

$\rho$ , respectively. Then, we will compute the minimum outage probability. Thus, we will derive the optimal save-ratio  $\rho$ .

In our simulation, we consider the BPSK modulation communication system. The channel is Rayleigh fading, and the path loss exponent  $\eta$  is set as 2. Without loss of generality, we assume that the source-to-receiver distance  $d_{s,d}$  is normalized to 1, and the channel between two nodes  $h_{s,d}$  is i.i.d. Rayleigh fading random variables with unit variance. Besides, we also assume the target transmission information  $Q = 1$ , and the time slot  $T = 1$ . Hence, the target transmission rate  $R \geq R_{\text{eff}} = Q/(1 - \rho)T$  bits/s.

Figure 1 shows the system mutual information versus the save-ratio. Figure 2 illustrates how the save-ratio  $\rho$  affects the outage probability for  $P_T = 1$ . As can be seen from Figure 2, the system mutual information  $I$  is increased with the save-ratio  $\rho$  changing from 0 to 1, due to the improvement of instantaneous SNR in the destination. The save-ratio  $\rho$  affects outage performance very significantly, which can be

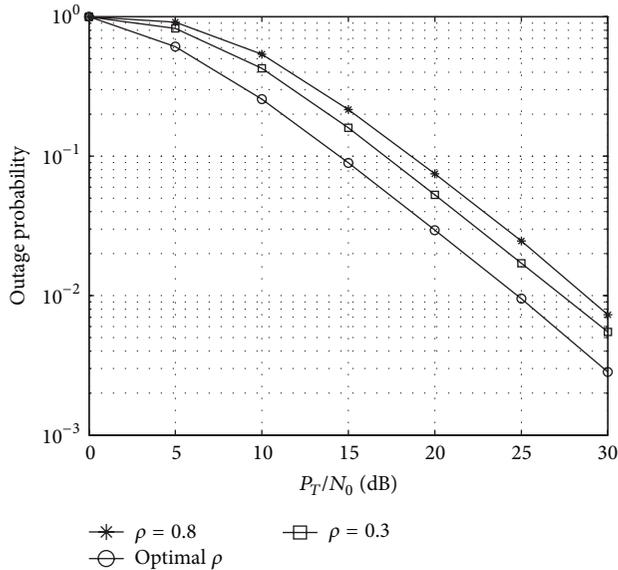


FIGURE 3: Outage probability comparison for optimal system and nonoptimal system.

seen from Figure 2. When the save-ratio is close to 0, almost no energy is needed for the transmitter to send the data, so the SNR is low, and the outage probability is almost 1. If the save-ratio is close to 1, though there is energy for the transmitter, the target transmission rate tends to  $+\infty$ . Hence, the outage probability is very high. As can be observed, there is an optimal decision on the save-ratio  $\rho$  of the transmitters to achieve the optimal outage performance.

In Figure 3, we plot the outage probability for the optimal system. From this figure, we see that the optimal save-ratio  $\rho$  scheme reduces the outage probability compared to the other cases.

## 6. Conclusions

In this paper, we studied a wireless system under energy harvesting conditions. We established the relationship between the system outage performance and the transmission protocol. Besides, we proposed an optimal method to optimize the system outage performance via finding the optimal save-ratio. We also characterized how the outage probability varies with the save-ratio and compared the outage performance between nonoptimal save-ratio and optimal save-ratio under the assumption of Rayleigh fading channel. Simulation results demonstrated that remarkable gain is achieved for the proposed optimal scheme. If we choose the optimal save-ratio for the transmission, we can achieve the outage performance minimization. And the optimal scheme can provide improved efficiency and reliability of the energy harvesting wireless sensor networks, which means such energy harvesting approaches are promising for further performance improvement.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This work is partially supported by the National Natural Science Foundation of China under Grants 61372088, 61271257, and 61328102 and the project funded by the Importation and Development of High-Caliber Talents Project of Beijing Municipal Institutions (CIT&TCD20130320).

## References

- [1] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Transactions on Wireless Communications*, vol. 12, no. 5, pp. 1989–2001, 2013.
- [2] Y. Chen, S. Zhang, S. Xu, and G. Y. Li, "Fundamental trade-offs on green wireless networks," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 30–37, 2011.
- [3] S. Yeh, "Green 4G communications: renewable-energy-based architectures and protocols," in *Proceedings of the Global Mobile Congress (GMC '10)*, pp. 1–5, IEEE, Shanghai, China, October 2010.
- [4] J. A. Paradiso and T. Starner, "Energy scavenging for mobile and wireless electronics," *IEEE Pervasive Computing*, vol. 4, no. 1, pp. 18–27, 2005.
- [5] A. M. Zungeru, L. M. Ang, S. Prabaharan, and K. P. Seng, "Radio frequency energy harvesting and management for wireless sensor networks," in *Green Mobile Devices and Networks: Energy Optimization and Scavenging Techniques*, pp. 341–368, CRC Press, 2012.
- [6] R. J. M. Vullers, R. van Schaijk, I. Doms, C. Van Hoof, and R. Mertens, "Micropower energy harvesting," *Solid-State Electronics*, vol. 53, no. 7, pp. 684–693, 2009.
- [7] T. B. Lim, N. M. Lee, and B. K. Poh, "Feasibility study on ambient RF energy harvesting for wireless sensor network," in *Proceedings of the IEEE MTT-S International Microwave Workshop Series on RF and Wireless Technologies for Biomedical and Healthcare Applications (IMWS-BIO '13)*, Singapore, December 2013.
- [8] S. H. Lee, R. Zhang, and K. B. Huang, "Opportunistic wireless energy harvesting in cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 9, pp. 4788–4799, 2013.
- [9] V. Liu, A. Parks, V. Talla, S. Gollakota, D. Wetherall, and J. R. Smith, "Ambient backscatter: wireless communication out of thin air," in *Proceedings of the ACM Conference on SIGCOMM (SIGCOMM '13)*, pp. 39–50, Hong Kong, China, August 2013.
- [10] O. Ozel and S. Ulukus, "Information-theoretic analysis of an energy harvesting communication system," in *Proceedings of the 21st IEEE International Symposium on Personal, Indoor and Mobile Radio Communications Workshops (PIMRC Workshops '10)*, pp. 330–335, IEEE, Istanbul, Turkey, September 2010.
- [11] J. Yang and S. Ulukus, "Optimal packet scheduling in an energy harvesting communication system," *IEEE Transactions on Communications*, vol. 60, no. 1, pp. 220–230, 2012.
- [12] Z. Wang, V. Aggarwal, and X. Wang, "Renewable energy scheduling for fading channels with maximum power constraint," in

*Proceedings of the 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton '13)*, pp. 1394–1400, Monticello, Ill, USA, October 2013.

- [13] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener, “Transmission with energy harvesting nodes in fading wireless channels: optimal policies,” *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 8, pp. 1732–1743, 2011.
- [14] Z. Wang and V. Aggarwal, “Power allocation for energy harvesting transmitter with causal information,” *IEEE Transactions on Communications*, vol. 62, no. 11, 2014.
- [15] G. E. Forsythe, M. A. Malcolm, and C. B. Moler, *Computer Methods for Mathematical Computations*, Prentice-Hall, 1976.

## Research Article

# Random Secure Comparator Selection Based Privacy-Preserving MAX/MIN Query Processing in Two-Tiered Sensor Networks

Hua Dai,<sup>1,2</sup> Tianyi Wei,<sup>3</sup> Yue Huang,<sup>1</sup> Jia Xu,<sup>1,2</sup> and Geng Yang<sup>1,2</sup>

<sup>1</sup>College of Computer Science & Technology, Nanjing University of Posts and Telecommunications, Nanjing 210023, China

<sup>2</sup>Jiangsu High Technology Research Key Lab for Wireless Sensor Networks, Nanjing 210003, China

<sup>3</sup>Bell Honors School, Nanjing University of Posts and Telecommunications, Nanjing 210023, China

Correspondence should be addressed to Hua Dai; [daihua@njupt.edu.cn](mailto:daihua@njupt.edu.cn)

Received 7 May 2015; Revised 4 August 2015; Accepted 24 August 2015

Academic Editor: Gwanggil Jeon

Copyright © 2016 Hua Dai et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Privacy-preserving data queries for wireless sensor networks (WSNs) have drawn much attention recently. This paper proposes a privacy-preserving MAX/MIN query processing approach based on random secure comparator selection in two-tiered sensor network, which is denoted by RSCS-PMQ. The secret comparison model is built on the basis of the secure comparator which is defined by 0-1 encoding and HMAC. And the minimal set of highest secure comparators generating algorithm MaxRSC is proposed, which is the key to realize RSCS-PMQ. In the data collection procedures, the sensor node randomly selects a generated secure comparator of the maximum data into ciphertext which is submitted to the nearby master node. In the query processing procedures, the master node utilizes the MaxRSC algorithm to determine the corresponding minimal set of candidate ciphertexts containing the query results and returns it to the base station. And the base station obtains the plaintext query result through decryption. The theoretical analysis and experimental result indicate that RSCS-PMQ can preserve the privacy of sensor data and query result from master nodes even if they are compromised, and it has a better performance on the network communication cost than the existing approaches.

## 1. Introduction

As wireless sensor networks (WSNs) have been widely used in a variety of important areas such as environment monitoring, medical care, national defense, and military, various security problems of data privacy are becoming more and more critical. For example, in the rare animal monitoring, the location of rare animals could be obtained for illegal hunting; in the application of smart home, the information for use of family hydroelectricity could be stolen for burglary. Therefore, privacy-preserving has become a very important issue in WSNs.

Most large-scale WSNs are expected to apply a two-tiered architecture with the resource-limit sensor nodes at the lower layer and resource-abundant master nodes at the upper layer, and this architecture is used to construct our concerned two-tiered wireless sensor networks (TWSNs) in this paper, as shown in Figure 1. The master nodes have abundant resources

of energy, computation, communication, and so forth, while the sensor nodes only have limited resources. The sensor nodes are only responsible of collecting data and periodically submitting it to a nearby master node for storage, which responds to the query requests from the base station (BS) and then returns the query results. Due to the simplicity of topological structure which contains multiple independent cells and the resource abundance of master nodes, TWSNs have a lot of advantages, such as stable link quality, simple route structure, and higher network scalability [1, 2].

However, because the master nodes are not only responsible of storing all the data from the sensor nodes, but also processing the query requests from BS, they are much more attractive and vulnerable to attackers in a hostile environment. Once a master node is compromised, serious threats could be brought to the data privacy of TWSNs. The attackers could utilize the compromised master nodes to obtain all the collected data of sensor nodes and the

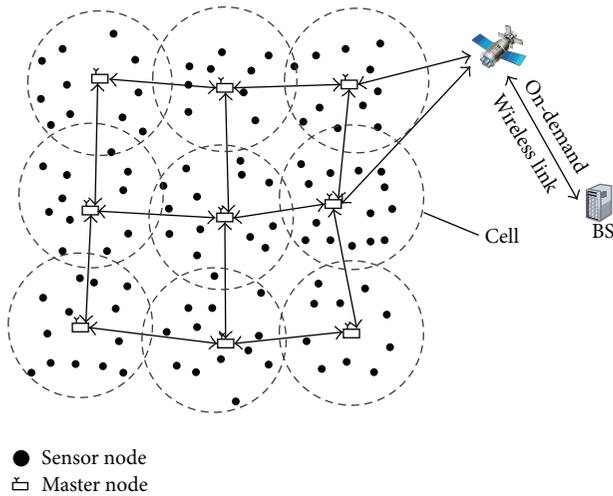


FIGURE 1: Architecture of TWSNs.

query results. Thus, it is necessary to investigate the privacy-preserving problems in TWSNs and develop efficient and effective solutions.

MAX/MIN query is a useful data query method to obtain the maximum or minimum data in the areas and epochs of interest. It can be utilized in event monitoring. For example, it can be applied to monitor the highest temperature in a warehouse so as to alarm the fire risk. The challenges to achieve privacy-preserving MAX/MIN query processing in TWSNs include the following:

- (i) How to make the master nodes realize secure comparisons of data items without knowing their real values and then determine the maximum or minimum value, that is, the query result.
- (ii) How to maximally reduce the communication cost of network, especially that of the sensor nodes due to their limited resources.

In this paper, we propose a privacy-preserving MAX/MIN query processing approach based on random secure comparator selection in TWSNs, which is denoted by RSCS-PMQ. The basic idea is as follows: once finishing data collection, the sensor nodes will encrypt the collected data into ciphertext and select the corresponding random secure comparators generated by using the 0-1 encoding [3] and hashed message authentication coding (HMAC) [4]. Then the ciphertext and the corresponding random secure comparators are submitted to the nearby master node. When the master node processes a query request from BS, it will utilize the algorithm MaxRSC to determine the minimal set of highest secure comparators, further determine the corresponding minimal set of candidate ciphertexts containing the query result, and return it to BS. After decrypting the received ciphertext, BS will obtain the query result in plaintext. Since the data storage and query response procedures in the master node do not involve the plaintext of collected data, the adversaries cannot read any hosted data or query result from the master node even if it is compromised. Thus, the

proposed RSCS-PMQ can achieve the privacy protection in MAX/MIN query processing. In addition, the evaluation indicates that RSCS-PMQ has better performance than other existing works on the network communication cost.

The main contributions of this paper are as follows:

- (i) We propose a comparison model based on random secure comparator selection through the 0-1 encoding and hashed message authentication coding, which supports data comparison without real values in master nodes.
- (ii) We design an algorithm to generate the minimal set of highest secure comparators based on the former comparison model, MaxRSC, which is the key algorithm to accomplish RSCS-PMQ.
- (iii) We provide the concrete protocols of achieving RSCS-PMQ, which consist of the data collection protocol and the query response protocol, and the latter protocol can protect data privacy from master nodes even if they are compromised.
- (iv) We analyze the privacy protection and communication cost of RSCS-PMQ and conduct performance evaluation through comprehensive simulation.

The rest of this paper is organized as follows. Section 2 gives an overview of related works. Section 3 describes related models and problem statement. In Section 4, we present the privacy-preserving MAX/MIN query protocols based on random secure comparator selection. Then, Section 5 analyzes the privacy and communication costs of our approach. We evaluate the performance through simulation in Section 6 and conclude this paper in Section 7.

## 2. Related Works

Data query is an important operation for events monitoring or data analysis in TWSN. The security issues are the hot spots in data query researches, such as privacy protection, integrity, or completeness verification. Recently, the secure range query [5–11] and top- $k$  query [12–17] have been broadly studied. However, there are limited researches on MAX/MIN query, and only [18, 19] propose solutions of privacy-preserving MAX/MIN query in TWSN.

Regarding the range query, a secure range query processing in TWSN is firstly proposed in [5], which employs bucket partition and symmetric encryption to achieve the privacy protection of collected data, and uses MAC to accomplish the completeness verification of query results. Based on [5], the spatiotemporal cross-check is introduced in [6, 7] to improve the efficiency of energy consumption. Furthermore, the spatiotemporal cross-check procedure of [6, 7] is optimized in [8], which balances the security and energy consumption and applies this method in multidimensional range query. A secure and energy-efficient range query processing protocol SafeQ is proposed in [9, 10], which is based on Prefix Membership Verification (PMV) [20] and neighborhood chain mechanism. In addition, the Bloom-Filter [21] is introduced to optimize the energy consumption. And a secure range

query protocol based on order-preserving function and link watermarking QuerySec is proposed in [11], which is capable of saving energy during query processing. However, these secure range query methods are not suitable for solving the privacy-preserving MAX/MIN query.

For top- $k$  query, the fine-grained verifiable top- $k$  query methods are proposed in [12, 13], whereby the network owner can verify the completeness and authenticity of query results in TWSNs. The verification code which embeds ordered and adjacent relationships of the collected data by HMAC is utilized in [14] to achieve the verifiable top- $k$  query processing. The symmetric encryption is applied in [15] to reduce the communication cost in verifiable top- $k$  query processing. Works in [12–15] merely support the completeness and authenticity verification of query results, but they cannot achieve privacy protection. To preserve privacy, the privacy-preserving top- $k$  query processing approaches based on Order Preserving Encryption [22] are proposed in [16, 17]. Though the top- $k$  query can be transformed into MAX/MIN query when  $k = 1$ , it is wasteful in energy consumption. The reason is that each sensor node should submit all the data collected in every epoch since top- $k$  is designed for obtaining the highest  $k$  data, where  $k$  is variable. In contrast, MAX/MIN query only has to submit the sole maximum or minimum value. Therefore, taking top- $k$  query as MAX/MIN query will result in high unnecessary data communication. In conclusion, the secure top- $k$  query methods are not suitable for solving the privacy-preserving MAX/MIN query.

For MAX/MIN query, the same PMV as in [9, 10], the symmetric encryption and HMAC are used in [18] to achieve the privacy-preserving MAX/MIN query processing. Since more codes are transmitted in data submission, which are generated by using the PMV and HMAC functions, the energy consumption of [18] is high, which will reduce the lifetime of whole network. In contrast, our former work [19] adapts 0-1 encoding verification instead of PMV to achieve an energy-efficient privacy-preserving MAX/MIN query which is denoted by EMQP. In EMQP, the codes generated by using 0-1 encoding and HMAC of same data are significantly less than those generated in [18], which can reduce the energy consumption of sensor nodes. Furthermore, in this paper, we will adopt random selection of codes on the basis of EMQP to save much more energy and accomplish better privacy-preserving MAX/MIN query processing.

There are also many secure aggregation methods, such as [23–26]. However, most of these works adapt the traditional multihop wireless sensor networks, and they are not suitable for TWSNs.

### 3. Models and Problem Statement

**3.1. Network Models.** We consider a similar TWSN model as in [5–17], as shown in Figure 1. The network is divided into multiple cells, each containing master node  $\mathcal{M}$  and several sensor nodes  $\{s_1, s_2, \dots, s_n\}$ , which is named after cell =  $(\mathcal{M}, \{s_1, s_2, \dots, s_n\})$ . In particular, the master nodes are powerful devices, which have abundant resources of energy, storage, and computation. Additionally, they are

also responsible for receiving and storing data collected by the sensor nodes and processing the query requests from BS, while the sensor nodes are cheap devices with limited resources. Each sensor node merely submits its collected data to the master node in the same cell. The master node can apply its long-distance and high-frequency communication capacity to communicate with the nearby nodes, which should then construct the upper-tier multihop networks. The query results will be returned from the queried master nodes to BS through the above networks. There is an on-demand wireless link (e.g., satellite) between the master nodes and BS to interact with each other. However, such wireless link is usually unstable and is of high consumption and low speed.

We assume that the time is divided into nonoverlapping epochs, and in every epoch  $t$ , the sensor node  $s_i$  collects  $V$  sensor data  $\{d_{i,1}, d_{i,2}, \dots, d_{i,V}\}$ . In TWSNs, BS owns the global network topological information, while a master node owns the network topological information of its located cell, and a sensor node only knows the locations of the master node in the same cell and 1-hop neighboring sensor nodes.

**3.2. MAX/MIN Query Models.** A MAX/MIN query in TWSN is a kind of query operation aimed at obtaining the maximum or minimum data among the data items collected in the specified epochs and area. Therefore, the following MAX/MIN query will be considered, which is denoted by a triple tuple:

$$Q = (\Theta, T, \Gamma), \quad (1)$$

where  $\Theta \in \{\text{MAX}, \text{MIN}\}$  refers to the query type,  $T$  is the set of queried epoch numbers, and  $\Gamma$  denotes the set of queried sensor nodes IDs which indicate a query region. For example, query  $Q = (\text{MAX}, t, \{s_1, s_4, s_6, s_{11}\})$  is aimed at getting the maximum data collected by sensor nodes  $s_1, s_4, s_6, s_{11}$  in the epoch  $t$ .

For simplicity, we focus on the simple MAX query aimed at one cell  $(\mathcal{M}, \{s_1, s_2, \dots, s_n\})$  and one epoch  $t$ ; that is  $Q = (\text{MAX}, t, \Gamma)$ , where,  $\Gamma \subseteq \{s_1, s_2, \dots, s_n\}$ . For other complicated queries covering multiple epochs and cells, it can be easily achieved by decomposing them into multiple simple ones. And we will conduct discussion in Section 4.5. Additionally, the MIN query is similar to the MAX query.

**3.3. Problem Statement.** In TWSNs,  $\mathcal{M}$  is too vulnerable and tends to be easily under attacks from adversaries, since they are not only responsible for storing all the data collected by the sensor nodes, but also responsible for processing the query requests from BS. If the collected data is in plaintext and  $\mathcal{M}$  is compromised, any data stored in  $\mathcal{M}$  and the query results generated by  $\mathcal{M}$  will be exposed to attackers, which tends to lead to privacy leakage. Therefore, it is necessary to take efficient and effective measures for privacy protection.

We adopt the same honest-but-curious threat model as in [27], where  $\mathcal{M}$  may try to breach privacy to steal sensitive data but faithfully obey protocols while processing the query requests. Additionally, BS and sensor nodes are also assumed to be credible in contrast to  $\mathcal{M}$ . Based on

the above assumption, for the sake of achieving privacy-preserving MAX/MIN query processing, the following conditions should be satisfied:

- (1) For the data collected by any sensor node in the network, only this sensor node and BS can obtain its real value in contrast to  $\mathcal{M}$ .
- (2) For the real value of query results, only BS can obtain it in contrast to  $\mathcal{M}$ .

Moreover,  $\mathcal{M}$  have abundant energy resources, while the sensor nodes are energy-limited, which results in the fact that the lifetime of the whole network totally relies on the energy consumption of sensor nodes. And most energy is consumed by communications according to [19]. Therefore, the communication cost of sensor nodes is a key metric for performance evaluation of query processing method in TWSN. We will conduct concrete evaluations of in-cell communication cost ( $C_s$ ) and query response communication cost ( $C_M$ ) in Section 6. The former represents the total energy consumption in bits incurred by data transmissions between the sensor nodes and  $\mathcal{M}$  per epoch, while the later refers to the total information in bits transmitted between  $\mathcal{M}$  and BS.

#### 4. MAX/MIN Query Processing with Random Secure Comparator Selection

We use the 0-1 encoding verification [3], which can be utilized to compare data items without knowing their values. Let  $x = b_1b_2 \cdots b_{w-1}b_w \in \{0,1\}^w$  be a binary string with  $w$  bits. The 0-encoding and 1-encoding of  $x$  are denoted by  $E_0(x) = \{b_1b_2 \cdots b_{i-1}1 \mid b_i = 0 \wedge 1 \leq i \leq w\}$  and  $E_1(x) = \{b_1b_2 \cdots b_i \mid b_i = 1 \wedge 1 \leq i \leq w\}$ , respectively, where  $|E_0(x)| + |E_1(x)| = w$ . For two data items  $x$  and  $y$ ,  $x > y$  if and only if  $E_1(x) \cap E_0(y) \neq \emptyset$ ; otherwise  $x \leq y$ . Obviously, if codes of  $x$  and  $y$  are of different types, they can be compared; otherwise they are incomparable.

In order to improve the efficiency of intersection computing, the numeralization functions are usually applied to convert the 0-1 encoding binary strings into numbers. Thus we adopt the same numeralization function  $\mathcal{N}(\cdot)$  as in [19], which satisfies the idea that, for any two 0-encoding or 1-encoding binary strings,  $a$  and  $b$ ,  $a = b$  if and only if  $\mathcal{N}(p) = \mathcal{N}(q)$ . Additionally, we utilize HMAC to realize one-wayness and collision resistance of encoding data. We denote HMAC function by  $H_g(\cdot)$ , where  $g$  is the secret key of HMAC, which is only shared in sensor nodes.

##### 4.1. Comparison Model Based on Random Secure Comparator Selection

**Definition 1.** For data  $x$ , after applying 0-1 encoding, numeralization, and HMAC, the two generated code sets are denoted by the secure comparators of  $x$ , where  $sc_0(x)$  and  $sc_1(x)$  are type 0 and type 1 secure comparators, respectively; that is,  $sc_0(x) = H_g(\mathcal{N}(E_0(x)))$ ,  $sc_1(x) = H_g(\mathcal{N}(E_1(x)))$ .

According to the data comparison property of 0-1 encoding verification, we do not have to compare any two data items

based on their values, but only the corresponding secure comparators. Thus, Lemma 2 is established.

**Lemma 2.** For data  $x$  and  $y$ , if  $sc_1(x) \cap sc_0(y) \neq \emptyset$ , then  $x > y$ ; otherwise  $x \leq y$ .

**Definition 3.** The random secure comparator of data  $x$  is denoted by  $\omega(x) = \text{rnd}\{sc_0(x), sc_1(x)\}$ , where  $\text{rnd}\{\cdot\}$  is the random selection function of a set. Its value is denoted by  $\omega(x).val$ ; that is,  $\omega(x).val = x$ , and its type is denoted by  $\omega(x).type$ , which is shown as follows:

$$\omega(x).type = \begin{cases} 0 & \omega(x) = sc_0(x) \\ 1 & \omega(x) = sc_1(x) \end{cases} \quad (2)$$

Given the random secure comparators of two different data items, we propose the *rsc\_compare* algorithm, as shown in Algorithm 1, to conduct secure comparisons according to Lemma 2 and Definition 3.

**Definition 4.** For two secure comparators,  $a$  and  $b$ , if and only if  $rsc\_compare(a, b) > 0$ , then  $a$  is higher than  $b$ , and we denote it by  $a > b$ , which means  $a.val \geq b.val$ ; if and only if  $rsc\_compare(a, b) < 0$ , then  $a$  is smaller than  $b$ , and we denote it by  $a < b$ , which means  $a.val < b.val$ ; otherwise,  $a$  and  $b$  are incomparable.

According to Definition 4, only if two secure comparators are of different types, they are comparable. It is remarkable that the relations  $>$  and  $<$  do not have transitivity.

**4.2. Algorithm for Generating Minimal Set of Highest Secure Comparators.** In this section, we firstly give the definition of the minimal set of highest secure comparators, which is the theoretical basis to achieve RSCS-PMQ. Then, we provide the generation algorithm of this set and analyze the probability of the amount of its elements.

**Definition 5.** Assume that  $S = \{\omega(x_1), \omega(x_2), \dots, \omega(x_n)\}$  is the corresponding set of random secure comparators of  $\{x_1, x_2, \dots, x_n\}$ . The minimal set of highest secure comparators is denoted by  $\Psi$ , where

- (1)  $\Psi \subseteq S$ ;
- (2)  $\forall \omega(x_i), \omega(x_j) \in \Psi \rightarrow \omega(x_i).type = \omega(x_j).type$ ;
- (3)  $\forall \omega(x_i) \in \Psi \wedge \omega(x_j) \in S - \Psi \rightarrow \omega(x_i) > \omega(x_j) \vee (\exists \omega(x_e) \in S - \Psi \wedge x_j \neq x_e \rightarrow \omega(x_i) > \omega(x_e) > \omega(x_j))$ .

According to Definition 5, the secure comparators in  $\Psi$  are of the same type; therefore, they are incomparable. Moreover, for secure comparators  $a \in \Psi$  and  $b \in S - \Psi$ , if they are of different types, the former is obviously larger than the latter; if they are of the same type, there must exist another secure comparator  $c \in S - \Psi$  with different type from  $a$  and  $b$ , which satisfies the idea that  $a$  is larger than  $b$ , while  $b$  is larger than  $c$ .

<b>Input:</b> secure comparators $a$ and $b$	
<b>Output:</b> if $a.val \geq b.val$ , 1 is returned, if $a.val \leq b.val$ , -1 is returned, and if $a.type = b.type$ , 0 is returned,	
<b>Procedures:</b>	
If $a.type = b.type$ , then return 0;	
If $a.type = 1$ , then	
If $a \cap b \neq \emptyset$ , then return 1;	//means $a.val > b.val$ ;
Else, return -1;	//means $a.val \leq b.val$ ;
Else	
If $a \cap b \neq \emptyset$ , then return -1;	//means $b.val > a.val$ ;
Else, return 1;	//means $b.val \leq a.val$ ;

ALGORITHM 1: *rsc\_compare*.

**Lemma 6.** Assume that  $\Psi$  is the minimal set of highest secure comparators of the data set  $\{x_1, x_2, \dots, x_n\}$ ; then we have

$$\max \{x_1, x_2, \dots, x_n\} \in \{a.val \mid a \in \Psi\}. \quad (3)$$

Lemma 6 can be easily deduced from Definition 5, which indicates that the maximum of  $\{x_1, x_2, \dots, x_n\}$  must exist in the corresponding data set of  $\Psi$ .

**Lemma 7.** Given data set  $\{x_1, x_2, \dots, x_n\}$ , its corresponding secure comparator set is  $S = \{\omega(x_1), \omega(x_2), \dots, \omega(x_n)\}$ ; then we have the probability of  $\Psi$  containing  $\lambda$  elements as follows:

$$\Pr(|\Psi| = \lambda) = \left(\frac{1}{2}\right)^\lambda, \quad (4)$$

where  $\lambda \in N^+$  and  $\lambda \leq n$ .

*Proof.* Assume that  $x_1 \geq x_2 \geq \dots \geq x_n$ , and  $\Psi$  has  $\lambda$  secure comparators; then  $\Psi = \{\omega(x_1), \omega(x_2), \dots, \omega(x_\lambda)\}$ . According to Definition 5, all the secure comparators in  $\Psi$  are of the same type. And  $\omega(x_{\lambda+1})$  must be of different type from the ones in  $\Psi$ ; otherwise  $\omega(x_{\lambda+1})$  also belongs to  $\Psi$ , which contradicts to the given assumption. Therefore, the probability of  $\Psi$  having  $\lambda$  secure comparators is equivalent to the probability that  $\omega(x_1), \omega(x_2), \dots, \omega(x_\lambda)$  are of the same type while differing from  $\omega(x_{\lambda+1})$ . Apparently, the probability of all  $\omega(x_1), \omega(x_2), \dots, \omega(x_\lambda)$  being type 0 and  $\omega(x_{\lambda+1})$  being type 1 is  $(1/2)^\lambda \cdot 1/2 = (1/2)^{\lambda+1}$ . Similarly, the probability is also  $(1/2)^{\lambda+1}$  under the reverse circumstance. Thus, the probability of  $\Psi$  having  $\lambda$  secure comparators is  $(1/2)^{\lambda+1} + (1/2)^{\lambda+1} = (1/2)^\lambda$ .  $\square$

The generation algorithm of the minimal set of highest secure comparators is given as Algorithm 2, denoted by MaxRSC.

As shown in algorithm MaxRSC,  $T_0$  and  $T_1$  are used to store the current sets of the highest and second highest secure comparators. The variable *flag* is to indicate the higher set between  $T_0$  and  $T_1$ , where *flag* = 0 indicates  $T_0$  is higher; otherwise  $T_1$  is higher. When the algorithm ends, the final minimal set of highest secure comparators is  $T_0$  if *flag* = 0; otherwise it is  $T_1$ . The algorithm is concise and direct, and its time complexity is only  $O(n)$ .

**4.3. Data Collection Protocol.** The data collection protocol is concerned with how a sensor node transmits its collected data items to  $\mathcal{M}$ . For each sensor node  $s_i$ , after collecting  $V$  data items  $\{d_{i,1}, d_{i,2}, \dots, d_{i,V}\}$  in epoch  $t$ , it performs the following procedure:

- (i) Determine the maximum data among the collected data items; that is,  $d_i = \max\{d_{i,1}, d_{i,2}, \dots, d_{i,V}\}$ .
- (ii) Compute the random secure comparator  $\omega(d_i)$ , and set its type according to the random selection.
- (iii) Encrypt  $d_i$  by using the key  $k_i$  shared with BS. The output ciphertext is denoted by  $(d_i)_{k_i}$ .
- (iv) Submit the following message to  $\mathcal{M}$ , where  $\text{id}(s_i)$  is the ID of  $s_i$ :

$$s_i \longrightarrow \mathcal{M} : \langle \text{id}(s_i), t, (d_i)_{k_i}, \omega(d_i) \rangle. \quad (5)$$

- (v) Once  $\mathcal{M}$  receives the above message from  $s_i$ , it will store the data of the message.

As shown in the above protocol, since the HMAC function is one-way and collision-resistant and sensor nodes only share the encryption and HMAC keys with BS, it is computationally infeasible to reveal the exact value to  $\mathcal{M}$ . Therefore, we can see that the data collection protocol can preserve data privacy from  $\mathcal{M}$ .

**4.4. Query Response Protocol.** The query response protocol is concerned with how  $\mathcal{M}$  cooperates with BS to accomplish the query requests from users. The main idea is that  $\mathcal{M}$  uses the MaxRSC algorithm to generate the minimal set of highest secure comparators on the basis of submitted random secure comparators of the queried sensor nodes and further determines the corresponding minimal set of candidate ciphertexts which is denoted by  $R$ . Then,  $\mathcal{M}$  returns it to BS. And the final query result will be determined after BS performs decryption on  $R$ . The concrete steps of query response protocol are as follows:

- (i) BS transmits the query request  $Q = (\text{MAX}, t, \Gamma)$  to  $\mathcal{M}$ , where  $\Gamma = \{s_1, s_2, \dots, s_m\}$ .
- (ii) Once  $\mathcal{M}$  receives the query request, it firstly loads the ciphertext  $(d_i)_{k_i}$  and the corresponding random

**Input:** The set of random secure comparators,  $S$   
**Output:** The minimal set of highest secure comparators,  $\Psi$   
**Procedures:**  
(1) Initialize  $T_0 = \emptyset, T_1 = \emptyset$  which are used to store type 0 and 1 secure comparators, respectively;  
(2) Fetch the first secure comparator from  $S$  into the variable  $a$ , and set the variable  $flag = a.type$ ;  
(3) If  $flag = 0$ , add  $a$  into  $T_0$ , otherwise add it into  $T_1$ ;  
(4) For each  $item \in S$ ,  
    If  $item.type = 0 \wedge flag = 0$ , then  
        If  $T_1 = \emptyset$ , then add  $item$  into  $T_0$ ;  
        Else  
            If  $\forall b \in T_1 (b < item)$ , then add  $item$  into  $T_0$ ;  
            Else, delete every  $b \in T_1$  if  $b < item$ ;  
    If  $item.type = 0 \wedge flag = 1$ , then  
        If  $\forall b \in T_1 (b > item)$ , then add  $item$  into  $T_0$ ;  
        Else if  $\forall b \in T_1 (b < item)$ ,  
            then clear  $T_0$ , add  $item$  into  $T_0$  and set  $flag = 0$ ;  
        Else, delete every  $b \in T_1$  if  $b < item$ ;  
    If  $item.type = 1 \wedge flag = 1$ , then  
        If  $T_0 = \emptyset$ , then add  $item$  into  $T_1$ ;  
        Else  
            If  $\forall b \in T_0 (b < item)$ , then add  $item$  into  $T_1$ ;  
            Else, delete every  $b \in T_0$  if  $b < item$ ;  
    If  $item.type = 1 \wedge flag = 0$ , then  
        If  $\forall b \in T_0 (b > item)$ , then add  $item$  into  $T_1$ ;  
        Else If  $\forall b \in T_0 (b < item)$ ,  
            then clear  $T_1$ , add  $item$  into  $T_1$  and set  $flag = 1$ ;  
        Else, delete every  $b \in T_0$  if  $b < item$ ;  
    End For  
(5) If  $flag = 0$ , then set  $\Psi = T_0$ ; otherwise, set  $\Psi = T_1$ ;

ALGORITHM 2: MaxRSC.

secure comparator  $\omega(d_i)$  received from each sensor node  $s_i \in \Gamma$  in epoch  $t$ . Assume that the loaded random secure comparators are  $\{\omega(d_1), \omega(d_2), \dots, \omega(d_m)\}$ . With them as the input, then  $\mathcal{M}$  generates the minimal set of highest secure comparators  $\Psi$  by using the MaxRSC algorithm, and the corresponding minimal set of candidate ciphertexts  $R$  is determined, where

$$R = \{(d_i)_{k_i} \mid \omega(d_i) \in \Psi\}. \quad (6)$$

(iii) Finally,  $\mathcal{M}$  constructs the following message and sends it to BS:

$$\mathcal{M} \longrightarrow \text{BS} : \langle \{\text{id}(s_i), (d_i)_{k_i} \mid (d_i)_{k_i} \in R\} \rangle. \quad (7)$$

(iv) When BS receives the response message from  $\mathcal{M}$ , it uses the key shared with the sensor nodes to decrypt the ciphertext in  $\mathfrak{R}$ , and then the final query result will be determined.

Similar to data collection protocol, it is also computationally infeasible for  $\mathcal{M}$  to get the real values and the query result in the query response protocol. Therefore, this protocol can also preserve data privacy from  $\mathcal{M}$ .

**Lemma 8.** For the determined minimal set of highest secure comparators  $\Psi$  and the corresponding minimal set of candidate

ciphertexts  $R$  in the query response protocol, we have the following:

- (1)  $|\Psi| = |R|$ , where  $|*|$  means the number of elements in the set  $*$ .
- (2) The query result of  $Q$  must be embedded in the ciphertext of  $R$ .

*Proof.* According to the construction of  $R$  in (6), we can easily have  $|\Psi| = |R|$ , and the values of secure comparators in  $\Psi$  are all embedded in the ciphertext of  $R$ . And Lemma 6 indicates that the maximum data among the data items collected by the queried sensor nodes (i.e., query result) must exist in the corresponding data set of  $\Psi$ . Therefore, the query result must be embedded in the ciphertext of  $R$  as well.  $\square$

**Lemma 9.** Assume that there are  $\lambda$  elements in  $R$ ; then we have its probability as follows:

$$\Pr(|R| = \lambda) = \left(\frac{1}{2}\right)^\lambda. \quad (8)$$

*Proof.* From Lemma 7, we can obtain the idea that the probability of  $\Psi$  containing  $\lambda$  secure comparators is  $(1/2)^\lambda$ . Additionally, Lemma 8 indicates  $|\Psi| = |R|$ . Thus, the probability of  $R$  having  $\lambda$  elements is also  $(1/2)^\lambda$ .  $\square$

According to Lemma 9, we can easily deduce Lemma 10.

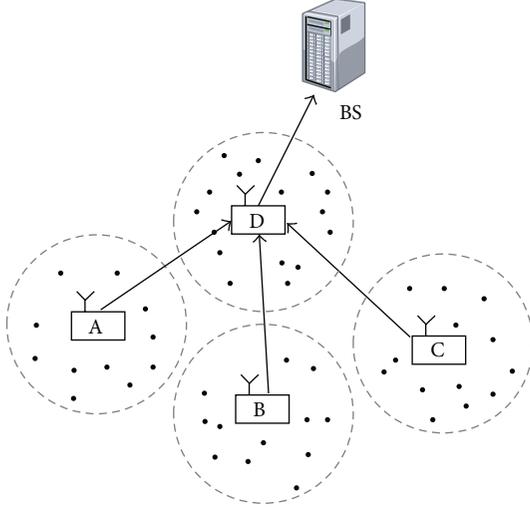


FIGURE 2: Complicated query example.

**Lemma 10.** *The mean quantity of elements in  $R$  is the mathematical expectation  $E(\lambda)$  of the ciphertext quantity in  $R$ , where*

$$\begin{aligned} E(\lambda) &= \sum_{\lambda=1}^n \lambda \cdot \Pr(\lambda) = \sum_{\lambda=1}^n \lambda \cdot \left(\frac{1}{2}\right)^\lambda \\ &= 2 - (n+2) \cdot \left(\frac{1}{2}\right)^n \end{aligned} \quad (9)$$

and  $E(\lambda) \approx 2$  when  $n$  is very large.

**4.5. Complicated Query Processing Method.** If complicated query  $Q$  involving multiple cells and epochs is applied, we can achieve it based on the basic ideas of the data collection and query response protocols. We give the overview of the complicated query processing method through an example.

As shown in Figure 2, there are four master nodes, A, B, C, and D, and BS composing the upper-tier tree-routing networks. Assume that the current query  $Q$  involves A, B, C, and D and several epochs. The main idea of processing  $Q$  is as follows. Firstly, A, B, C, and D use MaxRSC algorithm to determine their own minimal sets of highest secure comparators and the corresponding minimal set of candidate ciphertexts, which can be denoted by four pairs:  $(\Psi_A, R_A)$ ,  $(\Psi_B, R_B)$ ,  $(\Psi_C, R_C)$ , and  $(\Psi_D, R_D)$ , respectively. Then, A, B, and C submit  $(\Psi_A, R_A)$ ,  $(\Psi_B, R_B)$ , and  $(\Psi_C, R_C)$  to D on their own. And D takes  $\Psi_A$ ,  $\Psi_B$ ,  $\Psi_C$ , and its own  $\Psi_D$  as inputs into MaxRSC algorithm to determine the global minimal set of highest secure comparators and the global corresponding minimal set of candidate ciphertexts  $R$ . Obviously, the global query result is embedded in  $R$ , and then D submits  $R$  to BS. Consequently, BS decrypts the ciphertext in  $R$  and gets the final query result of the complicated query  $Q$ .

## 5. Protocol Analysis

### 5.1. Privacy-Preserving Analysis

- (1) Collected data privacy preservation: on the premise that BS and sensor nodes are credible in this paper,

the privacy of data collected by the sensor nodes can be preserved from  $\mathcal{M}$  only if it is ensured that it is impossible for  $\mathcal{M}$  to obtain the real value of any collected data. According to the data collection protocol, the data submitted by sensor nodes which are stored in  $\mathcal{M}$  are the ciphertext and HMAC codes instead of the plaintext. Since the HMAC algorithm is one-way and collision-resistant and the encryption and HMAC keys are only shared by the sensor nodes and BS, given a random secure comparator  $\omega(d_i)$  and ciphertext  $(d_i)_{k_i}$ , it is computationally infeasible for  $\mathcal{M}$  to obtain the value of the collected data  $d_i$ . And, for  $\mathcal{M}$ , the complexity to peek the privacy is equal to cracking the HMAC and encryption. Thus, our proposed RSCS-PMQ can protect the privacy of the collect data from master nodes.

- (2) Query result privacy preservation: as shown in the query response protocol,  $\mathcal{M}$  cooperates with BS to achieve the MAX/MIN query processing. During the procedure,  $\mathcal{M}$  takes the secure comparators as inputs to determine the minimal set of candidate ciphertexts embedding the plaintext query result through MaxRSC algorithm and then transmits it to BS. Consequently, BS decrypts the received ciphertext and gets the plaintext query result. Obviously,  $\mathcal{M}$  has no chance to touch any plaintext query result except for cracking encryption or HMAC. Therefore, RSCS-PMQ can protect the privacy of query results from the master nodes.

**5.2. Communication Cost Analysis.** To analyze the communication costs of data collection and query response protocols, we present the parameters as follows:

- $n$ : the number of sensor nodes.
- $l_{id}$ : the bit-length of a sensor node ID.
- $l_t$ : the bit-length of an epoch.
- $l_c$ : the bit-length of an encrypted data item.
- $l_h$ : the bit-length of a HMAC data item.
- $l_q$ : the bit-length of a query request.
- $w$ : the bit-length of a collected data item.
- $L$ : the average hops from a sensor node to  $\mathcal{M}$ .

According to the 0-1 encoding properties, there are  $w$  type 0 and type 1 secure comparators for every  $w$  bits data. Thus, the random secure comparator of a  $w$ -bits data item contains  $w/2$  HMAC data on average.

As shown in the data collection protocol, each sensor node submits a node ID, an epoch number, a ciphertext, and a secure comparator to  $\mathcal{M}$ . Therefore, the communication cost of data collection in the cell, denoted by in-cell communication cost ( $C_S$ ), can be calculated with

$$C_S = n \cdot \left( l_{id} + l_t + l_h \cdot \frac{w}{2} + l_c \right) \cdot L. \quad (10)$$

As shown in the query response protocol, the communication cost for executing a query consists of two parts:

TABLE 1: Computation cost analysis of PMV-PMQ, EMQP, and RSCS-PMQ.

Privacy-preserving MAX/MIN query methods	The quantity of operations in sensor nodes		The quantity of comparison operations in a storage node
	Encryption	HMAC	
PMV-PMQ	1	$[w + 2, 3w - 1]$	$(n - 1) * [w + 1, (2w - 2) * (w + 1)]$
EMQP	1	$w + 1$	$(n - 1) * [1, w * w]$
RSCS-PMQ	1	$[1, w]$	$(n - 1) * [1, w * w]$

Note:  $[a, b]$  is an interval range between the low-bound  $a$  and the upper-bound  $b$ .

one part is the communication cost of BS for sending query requests to  $\mathcal{M}$  and the other part is of  $\mathcal{M}$  for returning the feedback messages to BS. Additionally, Lemma 10 indicates that the mean quantity of ciphertext returned to BS is the mathematical expectation  $E(\lambda)$  of the ciphertext quantity in  $R$ . As a result, the calculation of query response communication cost is as shown which is denoted by  $C_M$ :

$$C_M = l_q + (l_{id} + l_e) \cdot E(\lambda). \quad (11)$$

According to Lemma 10, we have

$$C_M = l_q + (l_{id} + l_e) \cdot \left(2 - (n + 2) \cdot \left(\frac{1}{2}\right)^n\right) \approx l_q + 2(l_{id} + l_e), \quad (12)$$

when  $n$  is very large.

Then, we have the total communication cost  $C_{total}$  as follows:

$$C_{total} = C_S + C_M \approx n \cdot \left(l_{id} + l_t + l_h \cdot \frac{w}{2} + l_e\right) \cdot L + l_q + 2(l_{id} + l_e). \quad (13)$$

**5.3. Computation Cost Analysis.** We analyze the computation cost of proposed RSCS-PMQ and compare it with other privacy-preserving MAX/MIN query methods: PMV-PMQ [18] and EMQP [19]. First of all, since all of the three methods use the complex algorithms of encryption and HMAC in sensor nodes, the computation cost of sensor nodes is mainly caused by the encryption and HMAC. Secondly, the storage node  $\mathcal{M}$  determines the encrypted query results according to the intersections of paired code sets. To find out whether the intersection of two sets is null or not, many comparison operations are needed. As a result, the computation cost analysis of the three works is given in Table 1 on two aspects: the quantity of encryption and HMAC operations of a sensor node in an epoch and the quantity of comparison operations of  $\mathcal{M}$  in a query.

As shown in Table 1, PMV-PMQ, EMQP, and RSCS-PMQ perform the same quantity of encryption operations in a sensor node, but RSCS-PMQ performs less HMAC operations than the other two methods. RSCS-PMQ and EMQP perform the similar quantity of comparison operations, but they have general better performance than PMV-PMQ. Therefore, the RSCS-PMQ approach proposed by us is more efficient in computation cost than PMV-PMQ and EMQP.

We will not discuss the robustness of our method since it is not the focus of this paper. And we assume that the robustness is supported by the low-layer protocols.

## 6. Performance Evaluations

To analyze and compare the performance of protocols, we implement the proposed RSCS-PMQ, PMV-PMQ, and EMQP on the improved simulator of [28]. According to the experimental results of [19], we know that the energy consumed by data communication is much larger than that by the computation of encryption and HMAC. Therefore, this paper will focus on the evaluation of communication cost. We perform the evaluations on the following two aspects:

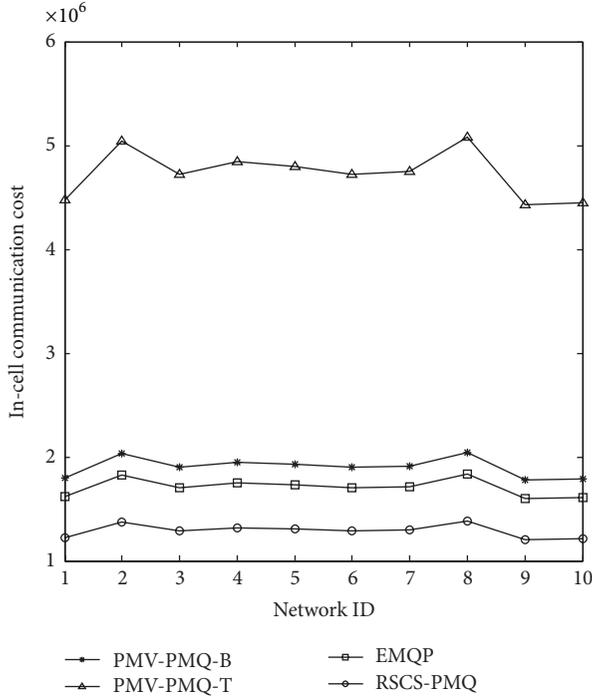
- (1) We firstly measure and analyze the in-cell communication cost ( $C_S$ ) of these three methods. Since the amount of codes for each collected data item constructed in PMV-PMQ is within a certain range, we consider the upper and lower bounds of PMV-PMQ in our evaluations, respectively, that is, the highest and lowest in-cell communication costs, which are denoted by PMV-PMQ-T and PMV-PMQ-B, respectively. Additionally, since the hash-based optimization in EMQP is also suitable for RSCS-PMQ and PMV-PMQ, which is aimed at reducing the length of HMAC data, this paper only compares the  $C_S$  of three methods without the hash-based optimization.
- (2) To evaluate the query response communication cost ( $C_M$ ) generated by  $\mathcal{M}$  and BS, we firstly measure the probability of  $R$  containing  $\lambda$  ciphertext and the average quantity of ciphertext in  $R$  in the RSCS-PMQ method. Then, we measure  $C_M$  of the three methods and calculate their proportions in the whole network communication costs while processing the MAX query.

The evaluations are performed on a PC with an Intel Core i5-3230M (quad-core 2.6 GHz) CPU and 8 G memory, running Windows 7 operating system, Eclipse, and Matlab. In addition, the experimental data set is randomly generated. In this simulation, the sensor nodes are assumed to be uniformly distributed in a cell covering a  $100 \times 100 \text{ m}^2$  area, and the communication radius of a sensor is 20 m. The default setting of other parameters is as shown in Table 2.

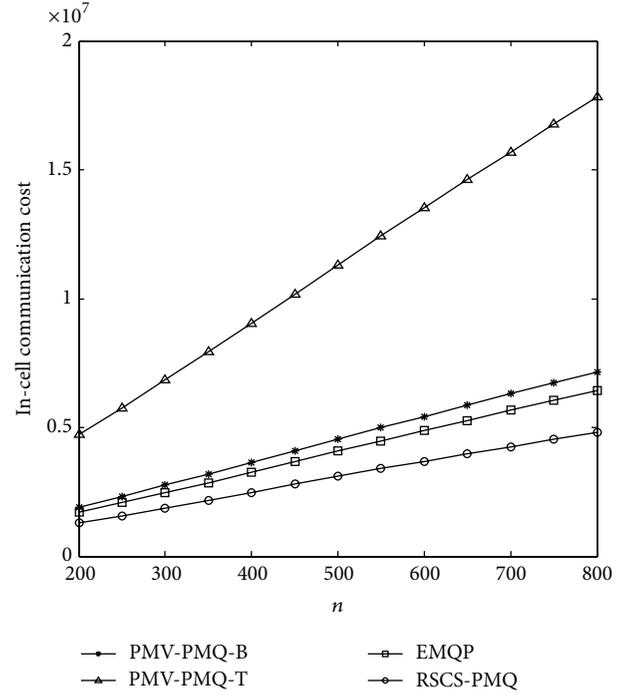
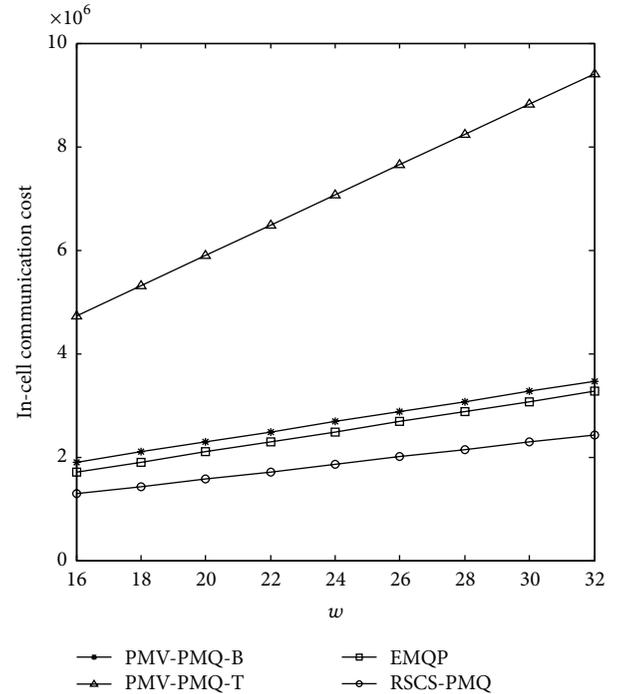
**6.1. In-Cell Communication Cost Evaluations.** In each measurement, we randomly distribute the sensor nodes and generate 10 networks with different topologies represented by different network IDs. Then, we can determine the communication cost of a MAX/MIN query by computing the average communication cost of these 10 networks.

TABLE 2: Default values of parameters.

Parameter	$l_{id}$	$l_t$	$n$	$w$	$l_c$	$l_h$	$l_q$
Value	32 bits	32 bits	200	16 bits	128 bits	128 bits	256 bits

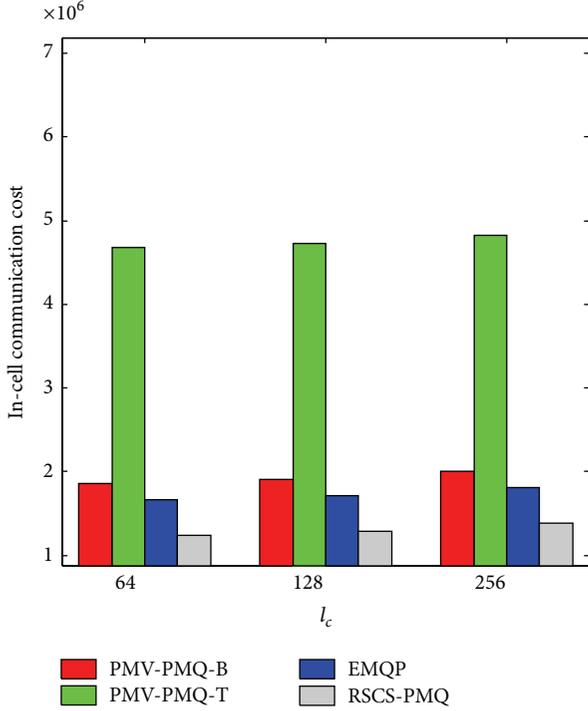
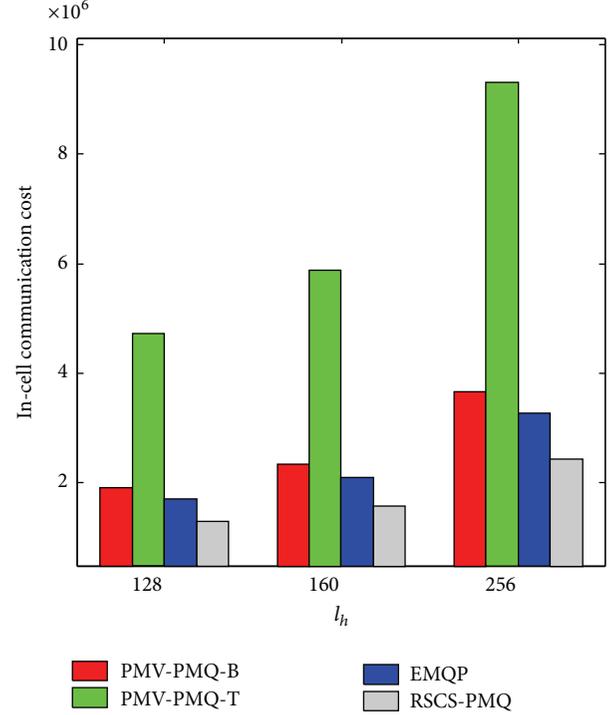
FIGURE 3:  $C_S$  versus network ID.

- (1)  $C_S$  versus network ID: Figure 3 shows that the  $C_S$  of RSCS-PMQ, EPRQ, and PMV-PMQ are all uniformly distributed in different topology networks. And the  $C_S$  of PMV-PMQ is the highest, and EPRQ has the mediate  $C_S$ , while RSCS-PMQ has the lowest. Under the experiment setting, the  $C_S$  of RSCS-PMQ is 32.23% lower than the lower bound of PMV-PMQ and 24.54% lower than the lower bound of EPRQ, since the amount of HMAC data used for secure comparison submitted from the sensor nodes to  $\mathcal{M}$  in the former method is smaller than that in the latter.
- (2)  $C_S$  versus  $n$  and  $w$ : as shown in Figure 4, when the amount of sensor nodes  $n$  increases, the  $C_S$  of RSCS-PMQ, EPRQ, and PMV-PMQ also increase, since the amounts of ciphertext and HMAC data transmitted in the network both increase. In accordance with Figure 5, we can see that the  $C_S$  of three methods also increase as  $w$  increases, because the amount of HMAC data used for secure comparison is in proportion to  $w$ . In addition, Figures 4 and 5 indicate that the  $C_S$  of three methods are in linear proportion to  $n$  and  $w$ , which is consistent with the theoretical analysis result shown in (10). Moreover, we have the idea that the  $C_S$  of RSCS-PMQ is significantly lower than that of EPRQ and PMV-PMQ, and the former is about 30%

FIGURE 4:  $C_S$  versus  $n$ .FIGURE 5:  $C_S$  versus  $w$ .

lower than the lower bound of PMV-PMQ and about 25% lower than the lower bound of EPRQ.

- (3)  $C_S$  versus  $l_c$  and  $l_h$ : we adopt different encryption and HMAC algorithms to set different  $l_c$  and  $l_h$ , respectively. For example,  $l_c$  could be 64, 128, and

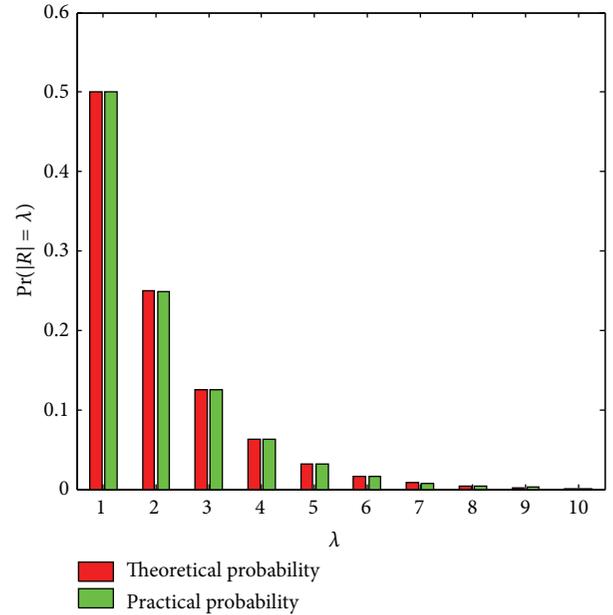
FIGURE 6:  $C_S$  versus  $l_c$ .FIGURE 7:  $C_S$  versus  $l_h$ .

256 bits if DES, IDEA, and AES-256 are adopted, respectively, while  $l_h$  could be 128, 160, and 256 bits if HMAC-MD5, HMAC-SHA1, and HMAC-SHA256 are adopted, respectively.

Figure 6 shows that the  $C_S$  of RSCS-PMQ, EPRQ, and PMV-PMQ have slow and unapparent increase as  $l_c$  increases, while they increase obviously as  $l_h$  increases. The reason is that there is only one encrypted data item in the message submitted from each sensor node to  $\mathcal{M}$ , but the amount of HMAC data is in proportion to the length of collected data  $w$ , which is obviously bigger than the former one. And the increasing of  $l_h$  has a more obvious influence on  $C_S$ . Similar to the results of evaluations (1) and (2) in this section, Figures 6 and 7 indicate that RSCS-PMQ is significantly lower than EPRQ and PMV-PMQ in  $C_S$ , and the former one is about 30% lower than the lower bound of PMV-PMQ and about 25% lower than the lower bound of EPRQ.

**6.2. Query Response Communication Cost Evaluations.** (1) Assume that the sensor nodes collect data in 10000 epochs and transmit the corresponding ciphertext and HMAC data to  $\mathcal{M}$ . And  $\mathcal{M}$  executes 10000 MAX queries aimed at each epoch mentioned above. We measure the probability and mean value of the amount of ciphertext in  $R$  returned from  $\mathcal{M}$  to BS. We repeat the experimental process 10 times and get the results as shown in Figures 8 and 9.

From Figure 8, we can see that the probability of  $R$  containing  $\lambda$  ciphertexts in the practical experiment totally corresponds to the theoretical probability computed with (8) in Lemma 9, which also proves the correctness of Lemma 9 from the point of experimental statistics. Additionally, based

FIGURE 8: The probability of  $|R| = \lambda$ .

on a large amount of experimental statistics, Figure 9 indicates that the mean quantity of ciphertext in  $R$  is in agreement with the mathematical expectation  $E(\lambda)$  computed with (9) in Lemma 10, and it is close to 2 as the amount of test samples becomes very large. The result verifies the correctness of Lemma 10 from the point of experimental statistics.

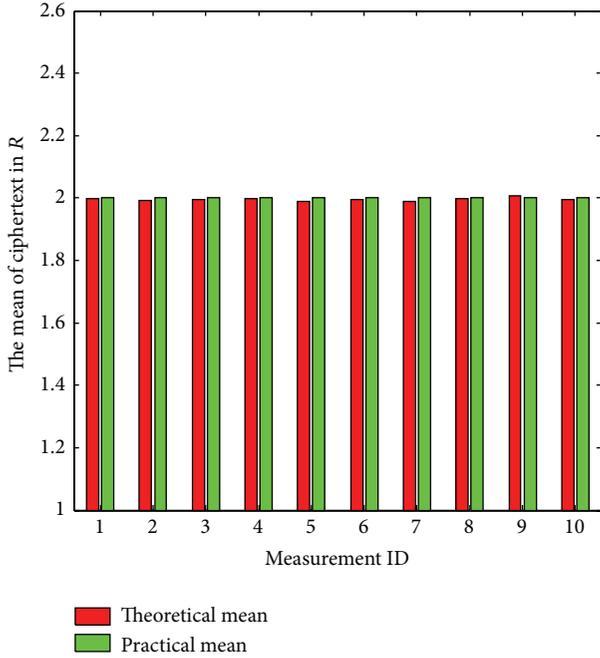
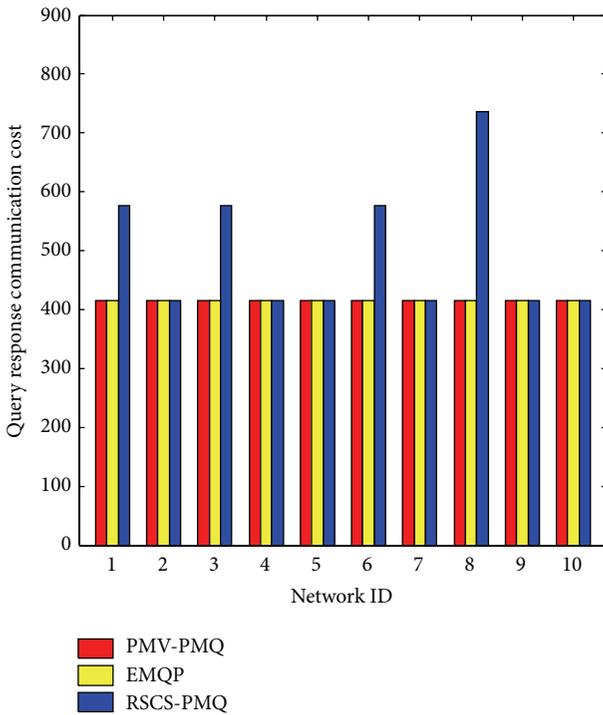
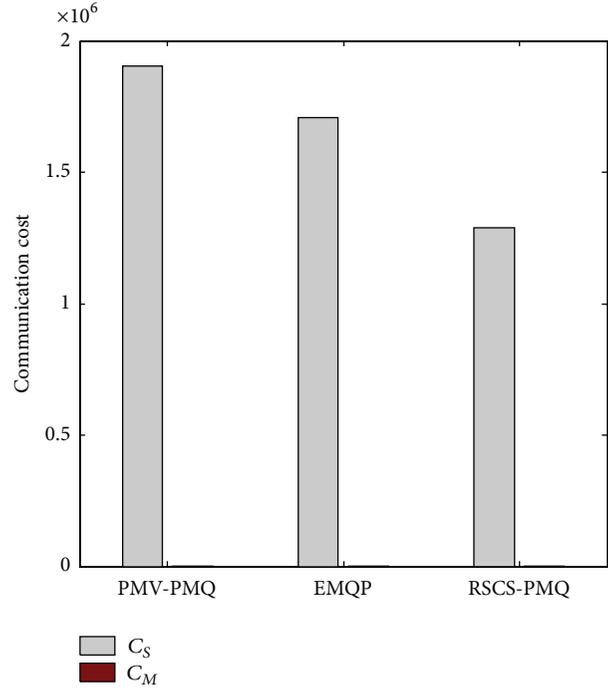


FIGURE 9: The mean of ciphertext in R.

FIGURE 10:  $C_M$  versus network ID.

(2) Based on the 10 groups of data transmitted from the sensor nodes  $\mathcal{M}$  under the 10 networks with random topologies in Section 6.1, we process 10 MAX queries, respectively. We test the query response communication costs ( $C_M$ ) and the average proportion of them in the total network communication costs ( $C_{\text{total}}$ ) for PMV-PMQ, EPRQ, and

FIGURE 11: The average of  $C_S$  and  $C_M$ .

RSCS-PMQ, respectively. The experimental results are shown in Figures 10 and 11.

Figure 10 indicates that the  $C_M$  of EPRQ and PMV-PMQ are constant and equal, while the  $C_M$  of RSCS-PMQ is about 20% higher than the former two methods. The reason is as follows:  $\mathcal{M}$  can only determine the ciphertext as the query result in EPRQ and PMV-PMQ, while the result returned by  $\mathcal{M}$  in RSCS-PMQ is the set  $R$  containing multiple candidate ciphertext. The probability statistics of the amount of ciphertext in  $R$  is as shown in Figure 8, and the mean quantity of  $R$  is about 2 according to Figure 9.

However, as shown in Figure 11, in the average  $C_{\text{total}}$  of PMV-PMQ, EPRQ, and RSCS-PMQ where  $C_{\text{total}} = C_S + C_M$ , the mean value of  $C_M$  is significantly smaller than that of  $C_S$ , and they merely account for a very small proportion of  $C_{\text{total}}$ , only 0.22%, 0.24%, and 0.38% on average, respectively. Here,  $C_S$  of PMV-PMQ is the lower bound of its in-cell communication cost. In addition,  $C_M$  is generated by the resource-abundant master nodes and BS. As a result,  $C_M$  has little impact on  $C_{\text{total}}$  which is mainly determined by  $C_S$  in contract, and  $C_{\text{total}}$  of RSCS-PMQ is lower than that of PMV-PMQ and EPRQ.

From the above experimental results and analyses, we can obtain the following: compared with the existing EPRQ and PMV-PMQ, the in-cell communication cost of RSCS-PMQ is lower, which is about 30% lower than the lower bound of EPRQ and about 25% lower than the lower bound of PMV-PMQ. Additionally, although the query response communication cost of RSCS-PMQ is higher than that EPRQ and PMV-PMQ, it only accounts for a very small proportion of the total network communication cost, lower than 1%, and so do the later methods. And the total communication cost of

RSCS-PMQ is lower than EPRQ and PMV-PMQ. Thus, the RSCS-PMQ proposed in this paper has a better performance than the existing works.

## 7. Conclusion

In this paper, we propose a novel random secure compactor selection scheme and a minimal set of highest secure comparators generating algorithm to achieve privacy-preserving MAX/MIN queries in two-tiered wireless sensor networks. Our technique can prevent the compromised master node from peeking at the hosted data and also ensure high query efficiency in network communication cost. Moreover, the efficacy and efficiency of our method are confirmed through detailed evaluations and analysis. In the future works, we will focus on the verification of query result completeness and further develop the key technique of this paper to support other types of data queries.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This research was supported by the National Natural Science Foundation of China under the Grants nos. 61300240, 61402014, 61572263, 61502251, 61472193, 61302157, 61373138, 61201163, and 61272084, the Natural Science Foundation of Jiangsu Province under the Grants nos. BK20151511 and BK20141429, the Project of Natural Science Research of Jiangsu University under Grants nos. 11KJA520002 and 14KJB520027, the Postdoctoral Science Foundation of China under the Grant no. 2013M541703, the Postdoctoral Science Foundation of Jiangsu Province under the Grant no. 1301042B, and Scientific & Technological Support Project (Society Development) of Lianyungang under the grant no. SH1306.

## References

- [1] O. Gnawali, K.-Y. Jang, J. Paek et al., "The tenet architecture for tiered sensor networks," in *Proceedings of the 4th International Conference on Embedded Networked Sensor Systems (SenSys '06)*, pp. 153–166, ACM, Boulder, Colo, USA, November 2006.
- [2] Y. Diao, D. Ganesan, G. Mathur, and P. Shenoy, "Rethinking data management for storage-centric sensor networks," in *Proceedings of the 3rd Biennial Conference on Innovative Data Systems Research (CIDR '07)*, pp. 22–31, Asilomar, Calif, USA, January 2007.
- [3] H.-Y. Lin and W.-G. Tzeng, "An efficient solution to the millionaires' problem based on homomorphic encryption," in *Applied Cryptography and Network Security: Third International Conference, ACNS 2005, New York, NY, USA, June 7–10, 2005. Proceedings*, vol. 3531 of *Lecture Notes in Computer Science*, pp. 456–466, Springer, Berlin, Germany, 2005.
- [4] H. Krawczyk, R. Canetti, and M. Bellare, "HMAC: keyed-hashing for message authentication," Tech. Rep. RFC 2104, Internet Society, Reston, Va, USA, 1997.
- [5] B. Sheng and Q. Li, "Verifiable privacy-preserving range query in two-tiered sensor networks," in *Proceedings of the 27th IEEE International Conference on Computer Communications (INFOCOM '08)*, pp. 46–50, Phoenix, Ariz, USA, 2008.
- [6] J. Shi, R. Zhang, and Y. C. Zhang, "Secure range queries in tiered sensor networks," in *Proceedings of the 28th Conference on Computer Communications (IEEE INFOCOM '09)*, pp. 945–953, Rio de Janeiro, Brazil, April 2009.
- [7] J. Shi, R. Zhang, and Y. Zhang, "A spatiotemporal approach for secure range queries in tiered sensor networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 1, pp. 264–273, 2011.
- [8] B. Sheng and Q. Li, "Verifiable privacy-preserving sensor network storage for range query," *IEEE Transactions on Mobile Computing*, vol. 10, no. 9, pp. 1312–1326, 2011.
- [9] F. Chen and A. X. Liu, "SafeQ: secure and efficient query processing in sensor networks," in *Proceedings of the 29th IEEE International Conference on Computer Communications (INFOCOM '10)*, pp. 1–9, IEEE, San Diego, Calif, USA, March 2010.
- [10] F. Chen and A. X. Liu, "Privacy and integrity-preserving range queries in sensor networks," *IEEE/ACM Transactions on Networking*, vol. 20, no. 6, pp. 1774–1787, 2012.
- [11] Y. Yi, R. Li, F. Chen, A. X. Liu, and Y. Lin, "A digital watermarking approach to secure and precise range query processing in sensor networks," in *Proceedings of the 32nd IEEE Conference on Computer Communications (INFOCOM '13)*, pp. 1950–1958, IEEE, Turin, Italy, April 2013.
- [12] R. Zhang, J. Shi, Y. Liu, and Y. Zhang, "Verifiable fine-grained top-k queries in tiered sensor networks," in *Proceedings of 29th IEEE International Conference on Computer Communications (INFOCOM '10)*, pp. 1199–1207, IEEE, San Diego, Calif, USA, March 2010.
- [13] R. Zhang, J. Shi, Y. Zhang, and X. Huang, "Secure top-k query processing in unattended tiered sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 9, pp. 4681–4693, 2014.
- [14] H. Dai, G. Yang, H. P. Huang et al., "Efficient verifiable top-k queries in two-tiered wireless sensor networks," *KSII Transactions on Internet and Information Systems*, vol. 9, no. 6, pp. 2111–2131, 2015.
- [15] X. Ma, H. Song, J. Wang, J. Gao, and G. Min, "A novel verification scheme for fine-grained top-k queries in two-tiered sensor networks," *Wireless Personal Communications*, vol. 75, no. 3, pp. 1809–1826, 2014.
- [16] C.-M. Yu, G.-K. Ni, I.-Y. Chen, E. Gelenbe, and S.-Y. Kuo, "Top-k query result completeness verification in tiered sensor networks," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 1, pp. 109–124, 2014.
- [17] X. Liao and J. Li, "Privacy-preserving and secure top-k query in two-tier wireless sensor network," in *Proceedings of the IEEE Global Communications Conference (GLOBECOM '12)*, pp. 335–341, Anaheim, Calif, USA, December 2012.
- [18] Y. Yao, N. Xiong, J. H. Park, L. Ma, and J. Liu, "Privacy-preserving max/min query in two-tiered wireless sensor networks," *Computers & Mathematics with Applications*, vol. 65, no. 9, pp. 1318–1325, 2013.
- [19] H. Dai, G. Yang, and X. Qin, "EMQP: an energy-efficient privacy-preserving MAX/MIN query processing in tiered wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 2013, Article ID 814892, 11 pages, 2013.

- [20] J. Cheng, H. Yang, S. H. Y. Wong, P. Zerfos, and S. Lu, "Design and implementation of cross-domain cooperative firewall," in *Proceedings of the 15th IEEE International Conference on Network Protocols (ICNP '07)*, pp. 284–293, Beijing, China, October 2007.
- [21] O. Rottenstreich and I. Keslassy, "The Bloom Paradox: when not to use a bloom filter," *IEEE/ACM Transactions on Networking*, vol. 23, no. 3, pp. 703–716, 2015.
- [22] R. Agrawal, J. Kiernan, R. Srikant, and Y. Xu, "Order preserving encryption for numeric data," in *Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD '04)*, pp. 563–574, Paris, France, June 2004.
- [23] M. Rezvani, A. Ignjatovic, E. Bertino, and S. Jha, "Secure data aggregation technique for wireless sensor networks in the presence of collusion attacks," *IEEE Transactions on Dependable and Secure Computing*, vol. 12, no. 1, pp. 98–110, 2015.
- [24] S. Roy, M. Conti, S. Setia, and S. Jajodia, "Secure data aggregation in wireless sensor networks: filtering out the attacker's impact," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 4, pp. 681–694, 2014.
- [25] Q. Zhou, G. Yang, and L. He, "A secure-enhanced data aggregation based on ECC in wireless sensor networks," *Sensors*, vol. 14, no. 4, pp. 6701–6721, 2014.
- [26] G. Yang, S. Li, X. Xu, H. Dai, and Z. Yang, "Precision-enhanced and encryption-mixed privacy-preserving data aggregation in wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 2013, Article ID 427275, 12 pages, 2013.
- [27] V. Bozovic, D. Socek, R. Steinwandt, and V. I. Villányi, "Multi-authority attribute-based encryption with honest-but-curious central authority," *International Journal of Computer Mathematics*, vol. 89, no. 3, pp. 268–283, 2012.
- [28] A. Coman, M. A. Nascimento, and J. Sander, "A framework for spatio-temporal query processing over wireless sensor networks," in *1st International Workshop on Data Management for Sensor Networks, DMSN '04, in Conjunction with VLDB 2004*, pp. 104–110, can, August 2004.