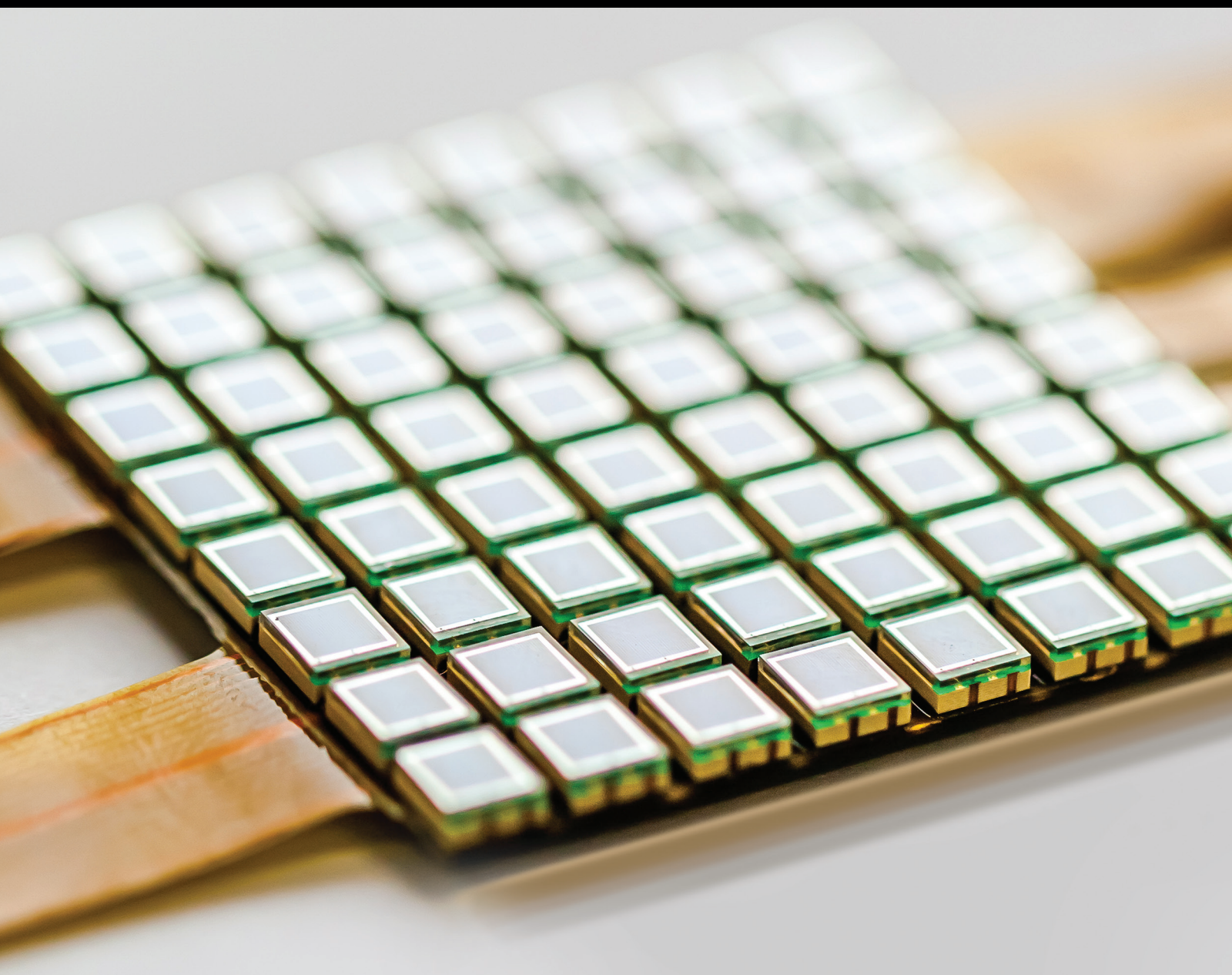
Journal of Sensors

Advanced Internet of Things and Big Data Technology for Smart Human-Care Services

Lead Guest Editor: Mucheol Kim
Guest Editors: Ka L. Man and Nurmamat Helil



Advanced Internet of Things and Big Data Technology for Smart Human-Care Services

Journal of Sensors

Advanced Internet of Things and Big Data Technology for Smart Human-Care Services

Lead Guest Editor: Mucheol Kim
Guest Editors: Ka L. Man and Nurmamat Helil

Editorial Board

Contents

Hindawi
Journal of Sensors
Volume 2019, Article ID 1654013, 3 pages
https://doi.org/10.1155/2019/1654013

Editorial

Advanced Internet of Things and Big Data Technology for Smart Human-Care Services

Mucheol Kim,[1] **Ka Lok Man**,[2] **and Nurmamat Helil**[3,4]

[1]*Chung-Ang University, Seoul, Republic of Korea*
[2]*Xi'an Jiaotong-Liverpool University, Suzhou, China*
[3]*University of Minnesota Twin Cities, Minneapolis, USA*
[4]*Xinjiang University, Xinjian, China*

Correspondence should be addressed to Mucheol Kim; mucheol.kim@gmail.com

Received 24 October 2018; Accepted 5 December 2018; Published 17 February 2019

Recently, the range of Internet of things (IoT) technology could be utilized in the real world by the rapid spread of sensor devices. Various sensed information such as temperature, illumination, pressure, and object detection has become a basic point for solving solutions to complex problems that have been difficult to solve by convergence with big data. Particularly, not only the detection of human health and changes in the natural environment but also the rapid development of the analysis and application technology is continuing. As a result, it is expected to dramatically improve the prediction of disaster and quality of life.

In order to implement IoT-based human-care technology for improving the quality of life, we aim to deal with sensor technology and communication technologies as well as big data analytics and its applications. This includes a wide range of application technologies that can be realized in real life as well as theoretical problem solving in relation to the IoT.

In this special issue, the latest research efforts on new theories, technologies, and research are related to IoT and big data technology for smart human-care services and to improve advanced technology. 25 research papers are published in this issue, covering a wide variety of topics including advanced sensor technology for smart human-care services, management issues of social network big data for smart human-care services, data and knowledge mining for smart human-care services, emerging IoT applications and systems for smart human-care services, and IoT solutions for well-being and active ageing. The highlights of these papers are summarized as follows.

The article titled "A Study on a Secure USB Mechanism That Prevents the Exposure of Authentication Information for Smart Human Care Services" by K. Lee et al. presents a secure USB mechanism that prevents leakages of authentication data and does not compare authentication data for smart human-care services, which have been a fundamental problem of existing flash drives.

The article titled "Fog Computing-Based IoT for Health Monitoring System" by A. Paul et al. was proposed to use fog computing to help monitor patients suffering from chronic diseases such that the data are collected and processed in an efficient manner.

The article titled "Beacon-Based Time-Spatial Recognition toward Automatic Daily Care Reporting for Nursing Homes" by T. Morita et al. presented an automatic daily report generation system that monitors the activities of nursing home residents. The proposed system estimated the multiple locations at which residents are situated with a BLE beacon and recognizes the activity of each resident from the estimated area information.

In the article titled "OperaBLE: An IoT-Based Wearable to Improve Efficiency and Smart Worker Care Services in Industry 4.0," L. Roda-Sanchez et al. introduced OperaBLE, a Bluetooth Low Energy (BLE) wearable proposal which is aimed at enhancing working conditions and efficiency in Industry 4.0 scenarios. The proposed approach had developed two algorithms for OperaBLE focused on power

awareness as Low-Frequency Movement Characterisation Algorithm (LoMoCA) and Adaptive Heart Rate Algorithm (AHRA).

In the article titled "Secure Data Encryption for Cloud-Based Human Care Services," T. Park et al. proposed efficient parallel implementations of Simeck family block ciphers on modern 64-bit Intel processors. In order to accelerate the performance, an adaptive encryption technique was also exploited for load balancing of the resulting big data. Finally, the proposed implementations achieved 3.5 cycles/byte and 4.6 cycles/byte for Simeck32/64 and Simeck64/128 encryption, respectively.

The article titled "Human Emotional Care Purposed Automatic Remote Portrait Drawing Generation and Display System Using Wearable Heart Rate Sensor and Smartphone Camera with Depth Perception" by G. Lee et al. proposed a system that automatically generates portrait drawings for the purpose of human emotional care. The user can recall the exciting and happy moment of the day through admiring the drawings and heal the emotion accordingly. To stylize photographs as portrait drawings, nonphotorealistic rendering (NPR) methods including a portrait etude stylization were proposed.

In the article titled "Biological-Signal-Based User-Interface System for Virtual-Reality Applications for Healthcare" by S. H. Nam et al., the proposed system extracted biological signal data from multiple biological signal data and simultaneously extracted and analyzed the data from a virtual reality-specific eye-tracking device that was developed so that users who develop healthcare contents based on virtual reality technology can easily use the biological signals.

In the article titled "Landmark-Guided Local Deep Neural Networks for Age and Gender Classification," Y. Zhang and T. Xu proposed to construct a local deep neural network for age and gender classification. The proposed model selected local image patches based on the detected facial landmarks.

The article titled "Video Retrieval System for Meniscal Surgery to Improve Health Care Services" by S. Amanat et al. proposed a training mechanism to the decumbent ratio through a video retrieval system. This research work was focused on developing a corpus and video retrieval system for meniscus surgery. Using the proposed system, surgeons can access guidance by watching the videos of surgeries performed by an expert and their seniors.

The article titled "Application of a Mobile Chronic Disease Health-Care System for Hypertension Based on Big Data Platforms" by D. Li et al. presented the provision of a practical assistant system for self-based patient health care, as well as the design of a complementary system for patient disease diagnosis.

The article titled "IoT Smart Home Adoption: The Importance of Proper Level Automation" by H. Yang et al. examined the smart home service features that current users require and empirically evaluates the relationship between the critical factors and the adoption behavior with 216 samples from Korea. The moderating effect of personal characteristics on behavior is also tested.

The article titled "Real-Time Cloud-Based Health Tracking and Monitoring System in Designed Boundary for Cardiology Patients" by A. Shahzad et al. was made to track and monitor the real-time medical information, bounded in an authorized area, through the modeling of private cloud computing. The private cloud-based environment is designed, for patient health monitoring called the bounded telemonitoring system, to acquire the real-time medical information of patients that resided in the boundary, inside medical wards and outside medical wards, of the medical center.

The article titled "IoT-Based Smart Building Environment Service for Occupants' Thermal Comfort" by H. Park and S.-B. Rhee presented an Internet of things (IoT) platform for a smart building which provides human-care services for occupants. The individual health profiles of the occupants were acquired by the IoT-based smart building, which uses the accumulated knowledge of the occupants to provide better services. The proposed model was based on the heat balance equation of human body and thermal characteristics of the occupants.

The article titled "Efficient Heterogeneous Network-Routing Method Based on Dynamic Control Middleware for Cyber-Physical System" by H. Lee et al. proposed a routing technique which enables network devices to communicate using different protocols. The proposed network-routing module can register devices with various protocols and improve the stability of the efficient heterogeneous network.

In the article titled "Intrusion Detection System Based on Evolving Rules for Wireless Sensor Networks," N. Lu et al. presented an evolving mechanism to extract the rules for intrusion detection. To extract diversified rules as well as control the quantity of rulesets, the extracted rules are examined according to the distance between the rules in the rule set of the same class and the rules in the rule set of different classes. Thereby, it alleviated the problem that the quantity of rules expands unexpectedly with the evolving genetic network programming.

The article titled "Multistandard Receiver Design for Telemedicine Monitoring System" by H. Wang et al. proposed a solution for aliasing in receivers that can reduce the limitations in sampling frequency, improve spectrum utilization, and realize multistandard receivers. This paper gave the constraint conditions for frequency selection and a method to process aliasing by more than two standard signals.

The article titled "T-S Fuzzy-Based Optimal Control for Minimally Invasive Robotic Surgery with Input Saturation" by F. Wang et al. presented a Takagi-Sugeno fuzzy model-based controller for a minimally invasive surgery robot with actuator saturation. The contractively invariant ellipsoid theorem was applied for the actuator saturation. The proposed approach can be derived using the infinity control theorem and parallel distributed compensation.

The article titled "Performance Analysis of Dual-Polarized Massive MIMO System with Human-Care IoT Devices for Cellular Networks" by J.-K. Hong archived the performance analysis when various human-care IoT devices

are connected to cellular networks via a dual-polarized massive MIMO system. There had not been an analysis that considers the performance between the massive MIMO system and IoT devices for cellular networks. In addition, dual-polarized transmit and receive antennas were proposed for BS and IoT devices, respectively, to install more antennas and achieve higher performances in limited space.

The article titled "A Development of Clinical Decision Support System for Video Head Impulse Test Based on Fuzzy Inference System" by D. T. A. Nguyen et al. presented the clinical decision support system for the video head impulse test (vHIT) based on a fuzzy inference system. It examined the eye and head movement recorded by the eye movement tracking device, calculates the vestibulo-ocular reflex (VOR) gain, and applies the fuzzy inference system to output the normality and artifact index of the test result.

In the article titled "Mcredit2: Enhanced High-Performance Xen Scheduler via Dynamic Weight Allocation," M. Kang and S. Lee proposed the Mcredit2 scheduler, which improves the Credit2 scheduler. The Credit2 scheduler takes no action when the load on a specific domain causes increased processor usage. The proposed Mcredit2 scheduler allowed a domain to quickly process loads by temporarily assigning a greater weight value to a host with high processor usage.

In the article titled "A Study of Prescriptive Analysis Framework for Human Care Services Based on CKAN Cloud," J. Gim et al. proposed a prescriptive analysis framework using a 5W1H method based on CKAN cloud. Through the CKAN cloud environment, IoT sensor data stored in individual CKANs can be integrated based on common concepts. As a result, it is possible to generate an integrated knowledge graph considering interoperability of data, and the underlying data is used as the base data for prescriptive analysis.

In the article titled "Theory and Application of Audio-Based Assessment of Cough," Y. Shi et al. introduced some successful cough monitoring equipment and their recognition algorithm in detail. It can be obtained that, firstly, acoustic variability of cough sounds within and between individuals makes it difficult to assess the intensity of coughing.

In the article titled "Analysis of Shooting Consistency in Archers: A Dynamic Time Warping Algorithm-Based Approach," C.-H. Quan et al. measured bow-forearm movement by using inertia sensors during archery shooting to evaluate the shooting consistency of archers. They also attempted to provide movement analysis tools that work precisely and conveniently. Shooting consistency was defined as the function of the dynamic time warping (DTW) distance between two time sequences of acceleration data calculated with the DTW algorithm.

In the article titled "The Intelligent Healthcare Data Management System Using Nanosensors," U.-O. Dorj et al. developed a design of the Intelligent Healthcare Data Management System (IHDMS) using nanosensors and composed an application for a mobile device. The proposed IHDMS could coordinate the healthcare data of the patients from nanosensors and transforms it into a worldwide consumed standard HL7 (Health Level Seven) for the conversion of healthcare data.

The article titled "Detection of Freezing of Gait Using Template-Matching-Based Approaches" by C. Xu et al. presented a template-matching-based improved subsequence dynamic time warping (IsDTW) method. The proposed approach could realize the real-time and high precision FOG detection and alarm.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

We would like to thank all the authors for their contributions and the anonymous reviewers for their voluntary support and constructive critiques in making this special issue possible.

Mucheol Kim
Ka Lok Man
Nurmamat Helil

Hindawi
Journal of Sensors
Volume 2018, Article ID 2089626, 17 pages
https://doi.org/10.1155/2018/2089626

Research Article

A Study on a Secure USB Mechanism That Prevents the Exposure of Authentication Information for Smart Human Care Services

Kyungroul Lee,[1] Insu Oh,[2] Yeunsu Lee,[2] Hyeji Lee,[2] Kangbin Yim,[2] and Jungtaek Seo[2]

[1]*R&DB Center for Security and Safety Industries (SSI), Soonchunhyang University, Asan 31538, Republic of Korea*
[2]*Dept. of Information Security Engineering, Soonchunhyang University, Asan 31538, Republic of Korea*

Correspondence should be addressed to Jungtaek Seo; seojt@sch.ac.kr

Received 31 January 2018; Accepted 27 August 2018; Published 29 October 2018

Academic Editor: Mucheol Kim

The secure USB flash drive was developed to improve the security of the conventional USB flash drive, which is vulnerable to leakages of internally stored data caused by extortion, loss, etc. However, it has been continuously reported that the secure USB flash drive, which protects data through the adoption of a wide range of security technologies in wide-ranging ways, cannot assure data security because of implementation and environmental vulnerabilities, eavesdropping, unlock commands, and reverse engineering. As such, there is growing demand for a more powerful secure USB flash drive to solve these fundamental problems. Therefore, this paper presents a secure USB mechanism that prevents leakages of authentication data and does not compare authentication data for smart human care services, which have been a fundamental problem of existing flash drives. The proposed mechanism provides better security than the existing secure USB flash drive by satisfying the need for confidentiality, integrity, authentication, and access control and safely protecting data from impersonation, man-in-the-middle, replay, and eavesdropping attacks by malicious attackers. An assessment of its security using the formalized verification tool AVISPA has proved that it is safe. Therefore, it is considered that a safer, more secure USB flash drive can be manufactured using the mechanism proposed in this paper.

1. Introduction

Generally speaking, a mobile storage unit USB flash drive is inserted into a USB port of a given host for use and is usually referred to as a USB memory stick or USB disk. The concept was originally introduced by an Israeli IT company in 2000 [1] and has been advanced significantly since then. Indeed, USB technology has been further developed because such devices offer easy connection and disconnection, free data change and deletion, fast data transfer, and high portability [2, 3]. In fact, these features of the USB flash drive have made it the most widely used storage unit currently in use [3].

The USB flash drive is mostly used for data storage and backup, booting disks, and portable program storage [3], and it must be capable of storing many different types of data in order to provide such functions. The various types of stored data include sensitive data such as public certificate, confidential business data, and personal information [3]. Therefore, a USB flash drive with guaranteed security is essential since serious damage can occur if such information is extorted by a malicious attacker.

Despite positive purposes, USB flash drives have problems regarding sensitive data leakage due to no additional security functions. Such a problem is the root cause of security threats that expose sensitive data through extortion and loss, and the reason for the threat is the fact that the data in the flash drive are stored in a raw form, rather than being encrypted in an altered form to cope with external attacks. Security threats by extortion enable a malicious attacker to extort a victim's private data stored inside a USB flash drive by seizing and inserting them into his own computer. There is a more serious problem in that there can be secondary or

tertiary victims if the extorted data are items of personal information such as public certificate or confidential business information. The security threat by loss refers to cases which a flash drive is lost as a result of a victim's error. As with the security threat by extortion, a third party who obtains a lost drive can take the data stored inside the drive by inserting it into his or her computer.

The secure USB flash drive was developed to solve the problem of data exposure by the security threats described above. A secure USB flash drive is a USB flash drive that comprises a security function designed to safely protect data stored inside a USB and supplements the vulnerability of a conventional USB flash drive with data encryption/decryption, user authentication and identification, prevention of arbitrary data copying, and data erasure technology to protect the data at the time of loss [3, 4]. There are many studies on diverse methods of applying these technologies to protect data, which can be broadly divided into the software type and the hardware type. The software type refers to methods of protecting internal data using only software and includes the data encryption/decryption method and the access control method [5]. The data encryption/decryption method uses a mathematical tool to encode and decode the data to prevent the inference of plain text, in order to solve the problem of data being stored in plain text in a conventional USB flash drive. Encryption/decryption methods include the encryption/decryption of a whole disk, the encryption/decryption of an image file to be used as a drive, and the encryption/decryption of a container file for stored files. Unlike the data encryption/decryption method, the access control method uses user authentication to block access by unauthorized users and allow access only by authorized users to the disk itself, image, or file used as a disk. Access control methods include access control [2] through user authentication [5] and access control by device authentication. On the other hand, the hardware method safely protects data using a separate hardware module. The main methods of protecting data include the method which uses an encryption module and the method which uses a flash drive controller [5]. The method using an encryption module uses a dedicated hardware security module to make it difficult to access secret information using reverse engineering, which is the fundamental problem of the software method. It includes the method of attaching a dedicated chip for user authentication and data encryption/decryption chip and a method of controlling access to the drive using a biometric authentication module to improve the security of user authentication. The method using a flash drive controller adds a security function to the existing flash drive controller to secure safety, instead of changing the hardware design so as to attach an encryption module. It includes the method of providing user authentication by adding a function for setting the password in the controller and the method of partitioning the flash drive into a general area and a secured area and performing access control and data encryption/decryption on the basis of the authentication data stored in the secured area, which cannot be accessed from outside [6].

Despite the protection of internally stored data using a secure USB flash drive which applies the above methods, there are still security problems related to data leakage to the outside due to such vulnerabilities as the leakage of a transferred password and authentication bypass or the exposure of the encryption/decryption key [2]. Such security threats to a secure USB flash drive are mainly classified into implementation and environmental vulnerabilities, eavesdropping, unlock commands, and reverse engineering. Implementation vulnerability refers to the design vulnerability that occurs when the security function of the management program that supports a secure USB flash drive has not been fully implemented. One example of this is the problem that occurs with the exception handling function when a public certificate is saved in a secure USB flash drive [3] and an unauthorized user accesses the data using the function that initializes the password and the data [7]. Moreover, another threat is caused by the incorrect implementation of data erasure by limiting the maximum password input count to protect the data at the time of loss. This enables an attacker to arbitrarily change the password input count limit and extort the password via a brute force attack. The environmental vulnerability is not the vulnerability existing in the secure USB flash drive but that which occurs because of the environment that connects the drive and the host or the host platform environment. For example, there is a potential threat of an unauthorized user accessing the important data stored in a secure USB flash drive when the secure USB is recognized by an installed VMware and directly accessed [3]. Other examples include the threat to the security function of a management program that is not operating properly following booting in the safe mode, forced termination of a management program, and the time difference during booting [3] and the threat of authentication bypass as a result of an exposed password or password hint by eavesdropping the data transferred between the secure USB flash drive and the host, because the security function was not considered when the USB interface was designed [6]. The vulnerability caused by an unlock function occurs by providing separate security domain connection and password initialization commands regardless of user authentication. Examples include the threat of accessing an inaccessible security domain by resending an unlock command or an unlock command with the authentication data [6]. The vulnerability caused by reverse engineering is an authentication bypass or encryption/decryption key exposure by the analysis of the security function provided by a management program [6].

There have been cases of internal data being exposed due to the failure of a secure USB flash drive to safely protect the data due to the various vulnerabilities described above, implementation vulnerability cases of abusing the initialization function, environmental vulnerability cases of eavesdropping, and unlock command vulnerability cases of password initialization. The implementation vulnerability case of using the initialization function includes the case of data not being deleted even after the initialization of the management program provided by the company "S" and thus the data being recovered with a data recovery tool in order to access the internally saved data without authentication [7]. The environmental vulnerability case of eavesdropping includes eavesdropping of plain text data transferred in

communication between the secure USB flash drive of companies A, S, L, and I and the host to capture the password or password hint used in user authentication [7]. The unlock command vulnerability case using password initialization includes normal login to a banking site using the reset password changed by sending the password reset command and the reset password to the secure USB flash drive for public certificate of company C. Such cases show that there is a serious risk of important data stored in secure USB flash drives being exposed and such simple data exposure can lead to serious monetary damages.

As such, it is urgently necessary to establish measures for manufacturing more powerful secure USB flash drives due to the failure of secure USB flash drives fitted with security technology to protect data. Therefore, this paper proposes a safe security USB mechanism that does not expose software-based authentication data in order to improve the security of existing secure USB flash drives. The fundamental problem with existing secure USB flash drives is that it is possible to access the data by bypassing authentication via exploiting the exposure of authentication data or modified authentication data stored within the drive and the existence of a routine that allows comparing authentications in the management program of the host. The mechanism proposed in this paper does not compare the authentication data and does not store the authentication data inside the drive. As such, it provides stronger security than existing secure USB flash drives by preventing data exposure through the authentication bypass caused by the abovementioned problems, thus satisfying the confidentiality, integrity, authentication, and access control requirements and safely protecting data from impersonation, man-in-the-middle, and eavesdropping attacks.

This paper is organized as follows: Section 2 reviews the classification, security requirement, and security technologies of the USB flash drive, which is the background of this study, and investigates vulnerabilities that cause data leakage in order to check the problems of existing secure USB flash drives, Section 3 describes the proposed mechanism and evaluates its security, and Section 4 presents the conclusion and outlines future studies.

2. Related Studies

The original secure USB flash drive was designed to safely store important information and assure the security of internally stored data at the time of loss, by applying hardware or software-based data encryption/decryption, user authentication and identification, prevention of arbitrary data copying, and data erasure technologies. However, despite the application of such wide-ranging security technologies, vulnerabilities due to the improper implementation of the security functions or the environmental limitations of the current platform have been reported and such vulnerabilities have led to such problems as stored data being leaked or extorted to the outside by a malicious attacker. Therefore, this study reviews the overview, classification, security requirement, and security technologies of existing USB flash drives as the background knowledge and investigates vulnerabilities that cause data leakage in order to identify the problems of existing secure USB flash drives.

2.1. Overview of Secure USB Flash Drives. A secure USB flash drive is a USB flash drive that protects the important data stored inside it using hardware or software technology [8]. Unlike conventional USB flash drives, the secure USB flash drive provides such security functions as user authentication and separates the flash drive into the general domain and secured domain for that purpose [9]. The secured domain stores special data such as authentication data for user authentication and the encryption/decryption key for data encryption/decryption. The general domain in which the user data are stored is safely protected by security technology such as user authentication and access control based on the special data. The security function through user authentication, for example, safely protects internally stored data by comparing the user authentication data stored in the secured domain with the requested authentication data to block access to the general domain by users who provide incorrect authentication data and allows access only by users who provide the correct authentication data. Such a security function is provided by attaching an additional hardware module that performs the security function or by applying software technology only. The method of attaching a hardware module uses the special protocol defined by the manufacturer to perform the security function, as the USB protocol failed to reflect the security requirement when it was first designed [3, 4]. The method of using software only runs special software implemented for security in the host to provide the security function. The various methods of safely protecting the data stored in a secure USB flash drive can be classified in the ways outlined below.

2.2. Classification of Secure USB Flash Drive Type. As described above, secure USB flash drives are classified into the hardware and software types. The hardware type of method uses a hardware module that provides an additional security function, such as user authentication or data encryption/decryption, while the software type of method uses the special software provided by the manufacturer as an additional security function, such as access control [5]. Each type is further classified in detail, as shown in Table 1.

2.2.1. Classification of Hardware Type. The hardware type adds a dedicated hardware module as the security function to safely protect the data and is divided into the controller method, which adds a security function such as user authentication in the controller of the existing USB flash drive and the encryption module method, which uses a dedicated encryption module to provide the security function, such as data encryption/decryption [10]. Table 2 shows the detailed classification of the hardware type.

The flash drive controller method assures the security by adding a security function to the existing flash drive controller instead of changing the hardware design by attaching a new encryption module. It is further divided into the controller-internal password method and the partitioning method. The controller-internal password method provides

Table 1: Classification of secure USB flash drive types.

Classification		Method	Applied security technology
Hardware	Controller method	Controller internal password authentication	User authentication
		Partitioning	Access control
	Encryption module method	Biometric authentication	User authentication
		Data encryption module	Data encryption/decryption
Software	Disk utilization method	Whole disk	Data encryption/decryption
		Image file	Data encryption/decryption
		Container	Data encryption/decryption
		Reserved domain	Access control
	Access control method	User authentication	Access control
		Device authentication	Access control

Table 2: Classification of hardware type.

Classification	Method	Description
Controller method	Controller internal password authentication	User authentication using password
	Partitioning	Access control using partitioning
Encryption module method	Biometric authentication	User authentication using biometric data such as fingerprint
	Data encryption module	Data encryption/decryption using a dedicated encryption chip

Table 3: Classification of software type.

Classification	Method	Description
Disk utilization method	Whole disk	Data protection by encrypting/decrypting whole disk
	Image file	Data protection by encrypting/decrypting the image file to be used as a disk
	Container	Data protection using container file
	Reserved domain	Use of reserved area as the secured domain
Access control method	User authentication	Access control through user authentication
	Device authentication	Access control through device authentication

user authentication by adding a function to set the password in the controller and protects the data by only allowing authenticated users to access the data. With the user authentication technology, the users register the password and are allowed to access internally stored data only when the password input during authentication matches the registered password [5]. The partitioning method divides the secure USB flash drive into partitions to add domains that provide the security function and control access through them. The flash drive is partitioned into the secured domain to store the user authentication data and the general domain to store the user data; access to the general domain is allowed only when the authentication data input during the authentication process matches the stored authentication data [6].

The encryption module method uses a dedicated hardware security module to make it difficult to access secret data through reverse engineering, which is the fundamental vulnerability of the software method. It is further divided into the biometric authentication method and the data encryption module method. The biometric authentication method attaches a biometric authentication module to the secure USB flash drive to protect the data by encrypting/decrypting the flash drive through user authentication [10]. In other words, it assures the security of data stored in the flash drive by preventing users who fail the biometric authentication from using the flash drive and allowing only users who pass the biometric authentication to use the flash drive. The encryption module method provides data confidentiality by attaching a dedicated encryption module which performs user authentication and data encryption/decryption. It prevents the leakage of original data by authenticating users with the encryption module and by encrypting/decrypting the data based on the encryption/decryption key generated inside the module for normal users and does not encrypt or decrypt the data otherwise. This method provides additional security by permanently deleting internally stored data when authentication fails a specific number of times or when the flash drive is disassembled by force [10].

2.2.2. Classification of Software Type. The software type safely protects the internally stored data using only software. It is divided into the disk utilization method using encrypting/decrypting [5] and the access control method through user authentication and device authentication [2]. Table 3 shows the detailed classification.

Table 4: Classification of security technology of secure USB flash drives.

User authentication technology	Data security technology		
User authentication and identification	Data encryption/decryption	Data erasure for protection after loss	Prevention of arbitrary data copying
Password	On-the-fly encryption	Tampering	Access control
Biometric authentication	Selective file encryption	Input count limitation	Management system

The disk utilization method is classified not according to the data protection type but rather according to the subject that is used as the disk. It is further divided into the method of using the whole disk, that of using an image file as a disk, that of using a container file for the data and security function, and that of using the reserved area of the system as a disk [10]. The method of using the whole disk protects the data by encrypting and decrypting the whole disk using the software provided by the manufacturer. It prevents the leakage of original data stored throughout the disk. More specifically, this method decrypts the whole disk for use as a conventional USB flash drive when an authorized user wants to use the disk and decrypts the whole disk to safely protect the internally stored data when the user terminates use of the disk [5]. The method of using an image file does not encrypt and decrypt the whole disk but instead generates an image to be used as a disk first and then encrypts and decrypts the generated image to protect the data. It not only prevents leaks of data stored in the drive but also quickly encrypts and decrypts the data since it does not process the whole disk [2, 10]. Although the detailed process of this method is the same as the method of using the whole disk, the difference here is that the subject is not a whole disk but an image to be used as a disk. The concept of the method of using a container file is similar to that of the hardware-based partitioning method, except that it uses the container file which partitions the domains with software instead of hardware partitioning. Thus, it safely protects data by using the container file to store the data needed for user authentication and encrypting and decrypting both the container file and data [5]. The method of using a reserved area stores secret data such as authentication data in a reserved space inside a USB flash drive and controls access to the stored data based on it. It prevents data access through an authentication bypass by concealing the authentication data or transformed authentication data in the reserved space [2, 10].

The access control method does not encrypt/decrypt the data stored inside a flash drive but allows access to the flash drive or an image to be used as a disk only by authorized users through user authentication and device authentication. Access control through user authentication controls access to internal data by authenticating users with the software provided by the manufacturer. It compares the registered authentication data and the requested authentication data and allows access only when they match and blocks access when they do not match, in order to assure safety [5]. Device authentication allows or blocks the use of a device based on unique device information such as the serial number. It prevents the leakage of confidential data by only allowing the use of an accepted device in the accepted space and by blocking the use of an unaccepted device in an accepted space or an accepted device in an unaccepted space based on the in/out policy. In addition, it can be combined with the access privilege of the user to prevent leakages of internal data by allowing only authorized users to use it, and even then only when the accepted device is connected in an accepted space, and by blocking its use if an accepted user does not have the privilege.

2.3. Security Technology of Secure USB Flash Drives. The secure USB flash drive is configured in the diverse ways described above, and each method safely protects the internal data by applying a wide range of security technologies. Such security technologies must satisfy the requirements of user authentication and identification, data encryption/decryption, and data erasure so as to protect the drive at the time of loss and also prevent arbitrary data copying [11]. Table 4 shows the leading security technologies designed for that purpose.

2.3.1. User Authentication Technology. User authentication technology identifies authorized users and allows only authenticated users to use services and is also known as electronic authentication technology in the computing environment. The leading electronic authentication technologies include knowledge authentication based on the user's memory, such as the password, possession authentication based on the medium possessed by the user, and biometric authentication based on the user's unique physical data [12]. User identification is checked using a wide range of factors. The most widely used electronic authentication technology is the password-based knowledge authentication technology, which performs registration and authentication based on the user-remembered password. Since this technology is dependent upon the user's memory, it does not require a separate medium and has the benefits of low cost and high user friendliness. Such benefits make it the preferred user authentication technology for secure USB flash drives, and most commercial secure USB flash drives authenticate users with the password. However, there is a possibility of the password being exposed since the passwords registered by authorized users are generally stored inside the flash drive; but a more serious problem is the fact that the routine procedure by which the registered password is compared with the input password exists on the host side, making it vulnerable to exposure of the password and bypassing of authentication through reverse engineering [6]. To solve the vulnerability of the password method, products that run the password-comparing routine in a separate module attached inside the flash drive have been developed. However, these products

also store the authentication data inside the flash drive and so have the fundamental problem of not being able to control access inside the flash drive and thus cannot prevent authentication bypass by password change and exposure.

To solve the fundamental problem of authentication data exposure, the method of attaching a biometric authentication module in the flash drive in order to activate the flash drive through biometric authentication has been developed. Fingerprint recognition technology is the most widely used biometric authentication method developed to supplement the problem with the password method. Users register their fingerprint through the fingerprint recognition module attached to the flash drive, and access to the flash drive is only allowed to the user whose inputted fingerprint matches the registered fingerprint. Despite this technological advance, the problem of extorting the fingerprint remains because authentication bypass through reverse engineering, like the password method, is not impossible although it is not easy [10].

2.3.2. Data Security Technology. The data stored inside a secure USB flash drive can be leaked if the flash drive is lost by extortion by a malicious attacker or by a user error if the flash drive is not equipped with a data security technology. For that reason, the data security technologies are applied to prevent data leakage to outside and uncontrolled data access. The various technologies include a data encryption/decryption technology designed to prevent leakages of original data, a technology for preventing arbitrary data copying to control access to the data, and a data erasure technology for protecting a flash drive from data leakage after a loss.

Data encryption/decryption technology generates a cryptogram from the plain text data using an encryption algorithm to protect the original data. The technology is divided into full disk encryption (FDE) and on-the-fly encryption (OTFE). The method of encrypting the whole disk encrypts and decrypts all data stored on the disk according to the user authentication result. All data are decrypted in order to use the flash disk as a conventional disk if the user authentication is successful, and the whole disk is encrypted again to prevent the leakage of original data after use [5]. However, as there is a possibility of data leakage as the data are saved in their original form after user authentication, the OTFE was developed to solve the problem by preventing data leakage even while the disk is in use. This method prevents the data leakage inherent to FDE since it only decrypts data to be used by the user, instead of the whole disk, and leaves unused data in the encrypted form even after successful user authentication [5].

Because of its innate characteristics, a USB flash drive is highly portable and there is a high risk of loss due to user error, and an attacker who possesses a lost drive can separate the memory storing the data and obtain the internally stored data using a specially produced tool. For that reason, an antitampering technology designed to detect malicious hardware manipulation of flash drives was developed to protect data in the event of loss of a flash drive. This technology permanently deletes the internally stored data to prevent data leakage if a malicious hardware disassembly or manipulation is detected. However, a user authentication bypass can occur through a brute force attack that enters all possible passwords to obtain the user passwords, instead of hardware manipulation. For that reason, a technology that limits the input count was developed to prevent brute force attacks. This technology permanently deletes the internally stored data to prevent data leakage if the specific input count is exceeded.

User authentication technology has a serious vulnerability in that authentication bypass can be achieved by extorting the stored authentication data or updating them with altered authentication data, since the authentication data are stored inside the flash drive. Since the root cause of this vulnerability is the accessibility to the data stored inside the flash drive, technologies for preventing the arbitrary copying of data have been developed and these can be divided into a technology for controlling access to data and a management system that allows or blocks use of the device. This access control technology validates the subjects accessing the data through user authentication and device authentication. If authentication fails, connection to the memory storing the data is blocked via a hardware method or the data are concealed by a software method of disabling data access. If authentication succeeds, the above methods are inactivated to allow data access. However, there is a problem in that original data can be acquired by bypassing authentication as the data are not encrypted for storage if only access control is provided [5]. On the other hand, the technology consisting of a management system that controls access to a device (instead of access to data) prevents the arbitrary copying of data by allowing or blocking the use of the device according to user authentication, user privilege, and the space where the device is used based on the established policy. In detail, the technology registers the user authentication data and privilege and the devices to be used and then allows the use of the device only to authorized users who pass the authentication, but only when the registered device matches. The use of the device is blocked to prevent the leakage of internal data if any of the registered data do not match. In addition, leakages of internal data can be prevented by allowing the use of devices only in registered safe areas, i.e., the internal network, and by blocking use from outside, in order to improve security [13].

2.4. Vulnerability of Secure USB Flash Drives. Even after applying the abovementioned security technologies to the secure USB flash drive, vulnerabilities related to external leakage of the data safely stored inside the flash drive have been detected. The main vulnerabilities include the implementation vulnerability, environmental vulnerability, unlock command, and reverse engineering [3]. Table 5 shows the vulnerabilities detected so far and the security technologies that have been disabled by the identified vulnerabilities.

The implementation vulnerability is a design vulnerability that occurs when the security function of the management program designed to support the secure USB flash drive is not properly implemented. Examples include public certificate bypass, password initialization, and input count manipulation [3, 7]. Public certificate bypass accesses the data using exception handling when the public certificate is stored in a secure USB flash drive. Specifically, since the

Table 5: Classification of vulnerabilities of secure USB flash drives.

Vulnerability	Classification	Description	Security technology
Implementation vulnerability	Exception of public certificate	Exception handling when the public certificate is stored to allow data access	(i) Prevention of arbitrary data copying
	Password initialization	Data access through recovery tool after password/data initialization	(ii) Data protection after loss
	Input count manipulation	Manipulation of password input count to infer password through brute force attack	(iii) Data protection after loss
Environmental vulnerability	VMware	Disabling of management program using VMware	(iv) User authentication and identification
	Direct memory access	Direct access to flash drive to read and write data	(v) Prevention of arbitrary data copying
	Safe mode	Booting in safe mode to disable the management program	(vi) User authentication and identification
	Forced termination	Termination and disabling of management program by force	(vii) Prevention of arbitrary data copying
	Booting time difference	Data access during the time the management program is not run in the booting process	(viii) Prevention of arbitrary data copying
	Eavesdropping	Analysis of data transferred between the host and the flash drive to obtain the password or password hint	(ix) Data encryption/decryption
Unlock commands	Secured domain access command	Sending of unlock command to access the secured domain	(x) User authentication and identification
	Command containing the authentication data	Sending of unlock command containing the authentication data to access the secured domain	(xi) User authentication and identification
Reverse engineering	Authentication bypass	Analysis of authentication process of management program to bypass authentication	(xii) User authentication and identification
	Exposure of encryption/ decryption key	Analysis of data encryption/decryption function of management program to obtain the encryption/decryption key	(xiii) User authentication and identification (xiv) Data encryption/decryption

public certificate stored inside a drive can be accessed without the approval of a secure USB flash drive, the internal data can be changed to .pfx, which is an extension of the public certificate, using this function, and then arbitrarily copied [3]. Password initialization allows an unauthorized attacker to access the data using the password and data initialization function. In detail, the attack uses the fact that the password initialization function, which deletes the internally stored data when the secure USB flash drive is reset to the initial condition, does not permanently delete the data. The data can be obtained using a recovery tool that restores the data [7]. Input count manipulation concerns the incorrect implementation of data erasure by limiting the maximum password input count to protect the data at the time of loss. It enables an attacker to arbitrarily change the password input count limit and extort the password through a brute force attack. In detail, it manipulates the password input count limit in the management program to infinity through reverse engineering and then tries all possible passwords to obtain the registered password.

Environmental vulnerabilities are not present in a secure USB flash drive itself but occur because of the environment between the drive and the host or the host environment. Examples include vulnerabilities of VMware, direct memory access, safe mode, forced booting, booting time difference, and eavesdropping. The vulnerability of VMware allows data access by recognizing the secure USB flash drive in the VMware, so that the management program providing the security function does not run. It can cause serious damage as it does not need to disable or bypass the security program and does not leave a usage record [3]. The vulnerability of direct memory access is provided by the operating system and allows an attacker to access the data by accessing the drive directly through implementation of or by using a tool instead of accessing data via the inside of the flash drive [3]. Other vulnerabilities that enable attackers to access internally stored data include booting in the safe mode or terminating the management program by force, so that the management program that provides the security function does not run, using the time before the security program is run during booting [3]. Lastly, the eavesdropping attack exploits the vulnerability caused by failure to consider the security function in the design of a USB interface so as to bypass authentication by obtaining the password and password hint transferred between the flash drive and the host, extorts the data by obtaining the transferred encryption/decryption key, and

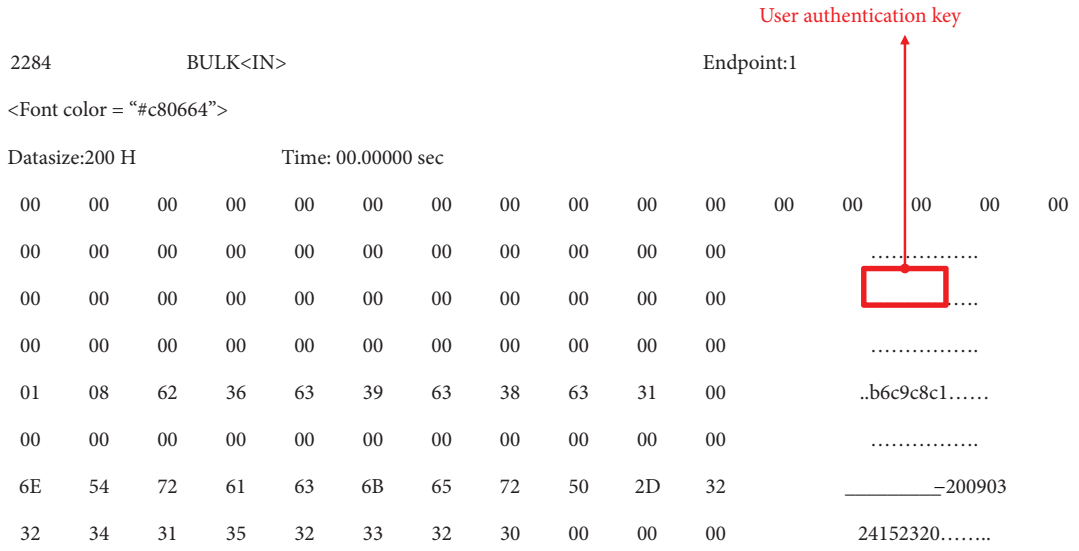


Figure 1: Example of authentication data exposure.

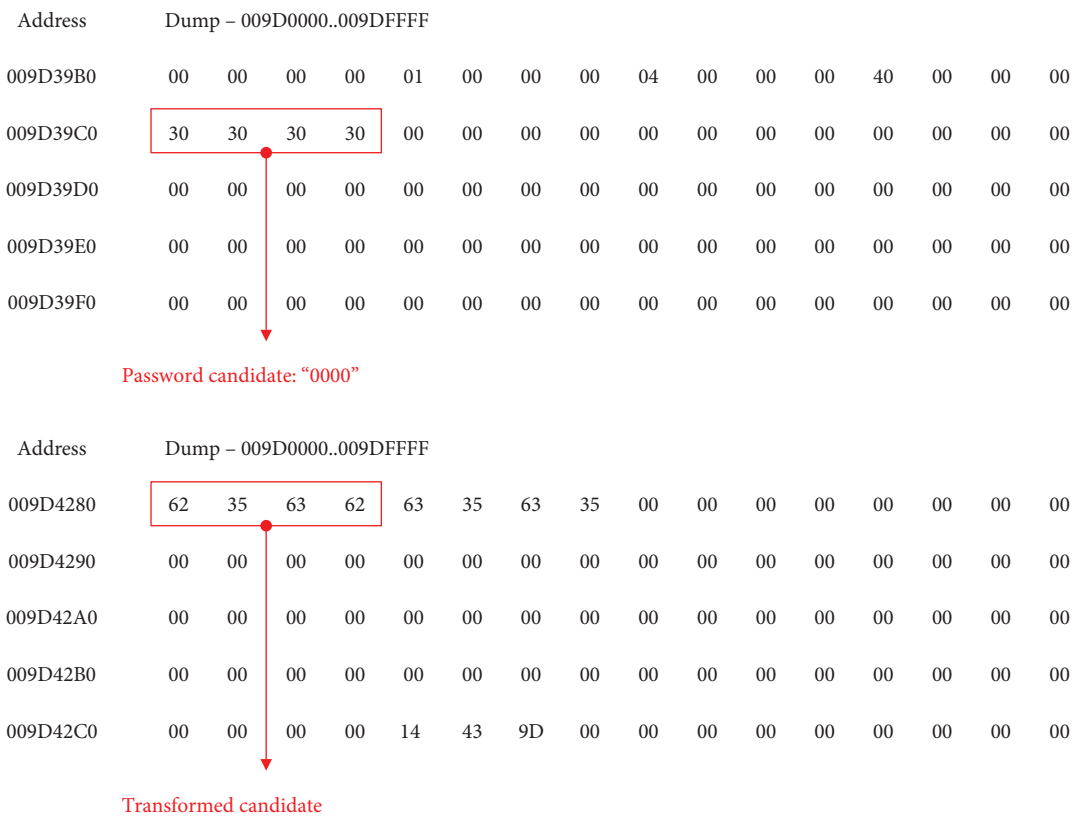


Figure 2: Example of use of exposed authentication data in management program.

then decrypts the encrypted data [6]. There have been actual cases of data being exposed in this way during communications, as shown in Figures 1 and 2.

The secure USB flash drive which separates the general domain and the secured domain in the hardware method has a special unlock command that connects the secured domain when user authentication is successful. In such a case, an attacker can analyze the unlock command and directly send a command enabling connection to the secured domain and extortion of internally stored data [6]. Such an unlock command can be configured by only the command code or additionally demanding the authentication code.

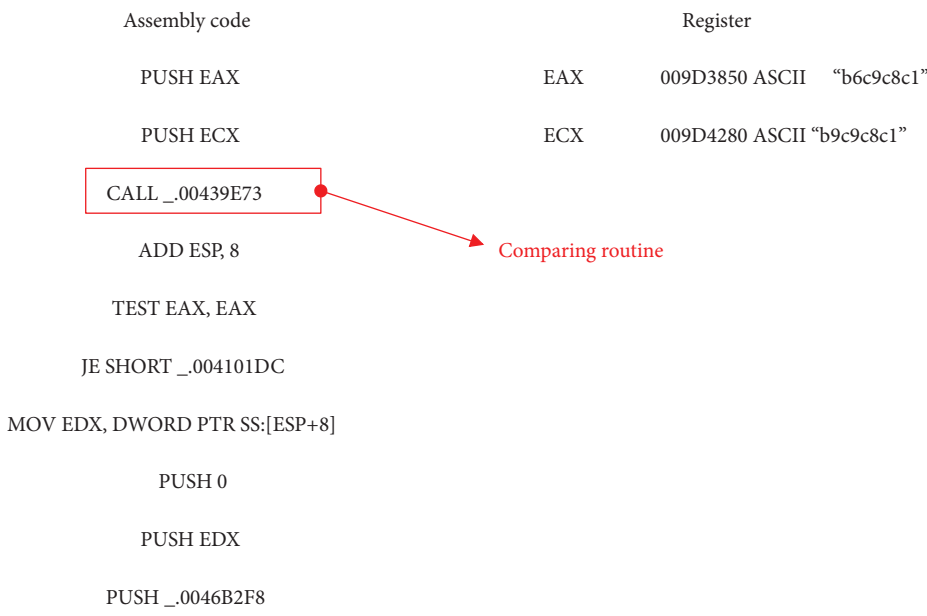


Figure 3: Example of comparison routine using exposed authentication data.

Although access to a secured domain is not possible if the correct authentication data are unknown, an additional command that transfers the authentication data is provided. As such, there have been cases in which the authentication data and the secured domain-connecting command were sent together to extort the internally stored data [8].

The vulnerability of reverse engineering allows an attacker to bypass user authentication and extort the exposed encryption/decryption key by analyzing the security function of the management program using reverse engineering. There have also been actual cases in which the routine that compares the authentication data and authentication bypassed by manipulating the routine in order to approve user authentication by force has been analyzed, as shown in Figures 3 and 4, or in which the encryption/decryption process has been analyzed, the exposed encryption/decryption key acquired, and the encrypted data decrypted [6].

Problems related to the failure of commercial secure USB products to safely protect internally stored data have been detected due to the various vulnerabilities described above. Therefore, this paper presents a mechanism that is designed to supplement the vulnerabilities of existing secure USB flash drives without exposing the authentication data in order to solve these problems.

3. Proposed Mechanism

Existing secure USB flash drives have problems in that the internally stored data can be extorted by exposing the password and bypassing user authentication due to the implementation vulnerabilities of public certificate exemption and password initialization; environmental vulnerabilities such as VMware vulnerability, direct memory access, and eavesdropping; and vulnerabilities using the unlock command. A more serious problem than any of the above vulnerabilities is that the current platform is based on the von Neumann architecture and loads the program code and the data needed for program operation in the memory. This allows program analysis through reverse engineering, which can lead to the exposure of key data. In other words, an attacker can analyze the internal operating process by reverse engineering the management program of a secure USB flash drive in order to extort the password and bypass authentication to access the data stored inside the drive. As an example, the management program run in a host performs the role of intermediary for delivering the authentication data between the secure flash drive and the user and storing the authentication data. Therefore, the key data must be allocated in the resource in which the management program is run. For that reason, the vulnerability which allows an attacker to access the resources containing the key data and extort the password or password-related authentication data and the architectural vulnerability which allows an attacker to bypass authentication, as the routine of comparing authentications exists in the management program of the host, have been identified. Since vulnerabilities with the flash drive due to the abovementioned fundamental reasons have been continuously discovered, studies are needed to remedy these problems.

Therefore, this paper proposes a safe software-based secure USB mechanism that disables authentication bypass while not exposing the authentication data in order to solve the aforementioned problems. The proposed mechanism uses a mathematical cryptologic tool to encrypt and decrypt the data and provides a function for authenticating users without exposing the authentication data, as well as preventing arbitrary copying of data, and erasing the data for protection after the loss of a flash drive. Table 6 shows the terms used in the proposed mechanism.

The proposed protocol consists of the registration process and the authentication process and provides disk management, user authentication, and data encryption/

1. Data encryption/decryption key is exposed

Address	Mem	Dump														
0012E978	4D	55	42	5A	46	50	51	4E	57	43	57	47	53	57	5A	47
0012E988	49	59	55	45	56	48	41	52	56	52	48	4C	48	45	48	59
0012E998	46	41	57	46	56	51	43	44	43	45	50	45	4a	50	59	46
0012E9A8	59	44	55	53	42	53	41	46	45	5f	4b	45	59	49	43	45

ASCII: MUBZFPQNWCWGSWZGIYUEVHARVRHLHEHY…

2. Attacker inputs a dummy password

Address	Mem	Dump														
0012D328	24	55	1b	0D	FA	93	7E	44	88	39	64	1F	A4	02	92	37
0012D338	C6	EA	D2	B9	C2	87	75	A3	04	23	D8	D2	71	6B	CF	FD
0012D348	A3	B3	D5	9A	24	32	38	6B	DD	AC	2D	65	4D	D9	53	F7
0012D358	7B	6A	67	1A	B2	F1	DD	A0	3F	0A	78	A6	C4	A5	C0	D3

3. Attacker replaces the key with the exposed one

Address	Mem	Dump														
0012D328	24	55	1B	0D	EA	93	7E	44	88	39	64	1F	24	02	92	37
0012D338	C6	EA	D2	B9	CB	87	75	A3	04	23	D8	31	71	6B	CF	FD
0012D348	A3	B3	D5	9A	24	32	38	6B	DD	AC	2D	65	4D	D9	53	F7
0012D358	7B	6A	67	1A	B2	F1	DD	A0	3A	0A	78	A6	C4	A5	C0	D3

Address	Mem	Dump														
0012E978	4D	55	42	5A	46	50	51	4E	57	43	57	47	53	57	5A	47
0012E988	49	59	55	45	56	48	41	52	56	52	48	4C	48	45	48	59
0012E998	46	41	57	46	56	51	43	44	43	45	4D	45	4a	50	59	46
0012E9A8	59	44	55	53	42	53	41	46	45	5F	4b	45	59	49	43	45

4. Authentication is by passed

Figure 4: Example of encryption/decryption key exposure.

decryption functions to assure data security. Such functions guarantee the security of the proposed mechanism by satisfying the requirement for confidentiality, integrity, authentication, and access control and by safely protecting the data from impersonation, man-in-the-middle, and eavesdropping attacks.

3.1. Registration Process. The registration process of the proposed mechanism consists of the following two steps: The first step is to register the users and encrypt the DIF, which is generated with the user-input authentication data and disk data. Specifically, the management program is run for registration and authentication in the host and the management program generates a DIF based on the authentication data received from a user and the data needed to generate the DIF. The user ID, which is needed for user authentication in the second step, is inserted in a reserved space of the file system header in the generated image file, and the generated image file is encrypted using the hash value of the user PW as the key, thus completing the first step of user registration. The second step authenticates a registered user by decrypting the encrypted DIF with the user-input authentication data and compares it with the registered authentication data to authenticate the user. The key to the user authentication process is to compare the ID inserted in the file system header during the registration process and the ID received from the authentication process to validate the user. The process decrypts the file system header encrypted with the hash value of the PW received in the authentication process and compares the received ID and the decrypted ID. The user is validated and registration is completed when the IDs match. Figure 5 shows the proposed registration process.

Table 6: Terms.

Term		Description
Status information	U	User
	ID	ID (identifier)
	ID'	ID rereceived by a user for disk decryption
	PW	Password
	PW'	PW rereceived by a user for disk decryption
	DIF	Disk image file
	ENDIF	Encrypted disk image file
	DEDIF	Decrypted DIF
Operating status	F	Format
	RUR	Request user registration
	RAI	Request authentication information
	TAI	Transfer authentication information
	IU	Identify user
	RGDIF	Request generation of DIF
	TDI	Transfer disk information
	GDIF	Generate DIF
	IID	Insert ID in an empty space in the file system header
	AU	Authenticate user
	FR	Finish registration
	MD	Mount disk
	SDIF	Select DIF
Encryption/ decryption	$h()$	Hash operation
	H	h(PW), result of hash operation based on PW
	H'	h(PW'), result of hash operation based on PW'
	$C = E_K(P)$	Result C of encryption of the plain text P based on the key K
	$P = D_K(C)$	Result P of decryption of the cyphered data C based on the key K
	EDIF	Encrypt DIF
	RDDIF	Request decryption of DIF
	RDFSH	Request decryption of file system header
	DFSH	Decrypt file system header
	DDIF	Decrypt DIF
Entity	User	User authentication
	Management program	Management program
	USB flash drive	USB flash drive

Step 1. A user requests registration with the management program when using the USB flash drive and then the management program confirms it (RUR).

Step 2. The management program requests the authentication data such as ID and PW, from the user requesting registration (RAI).

Step 3. The user inputs and sends the memorized ID and PW requested by the management program (TAI).

Step 4. The management program checks if the ID received from the user is in its database and approves the user's registration if it does not exist in the database (IU). Step 2 is repeated if the received ID exists in the database.

Step 5. If the user registration is approved, the management program requests generation of an image file to be used as a disk (RGDIF) and demands the data, such as the disk name and size of the DIF to be generated.

Step 6. The user inputs the DIF and the related data requested by the management program and sends them to the management program (TDI).

Step 7. The management program generates a file to be used as the disk inside the USB flash drive based on the disk image file and the related data received from the user (GDIF).

Step 8. The disk is formatted to add the data for disk recognition after the file to be used as a disk has been generated in the USB flash drive (F).

Step 9. The proposed mechanism uses the ID as the data for authentication in order to assure that the authentication data are not exposed and the ID received from the user (ID') is inserted into the reserved space of the header, such as the boot record of the file system (IID). The inserted ID is used to verify the user afterward. In detail, the inserted ID is encrypted with the hash value of the PW in the registration process and is decrypted with the hash value of the PW' received from the user in the later authentication process in Step 15. The user is authenticated in Step 15 by comparing the user-input ID' with the currently inserted ID.

Step 10. The management program performs the hash operation of the PW received from the user to generate the key for encryption of the image file to be used as a disk ($H = h(PW)$). The hash operation is necessary since using the PW directly as the key for encryption/decryption has the problem of PW leakage, because the time the PW is exposed outside increases during the encryption/decryption process. Performing the one-way hash operation prevents inference of the PW even when a third party obtains the hash value.

Step 11. The management program begins encryption of the DIF in the flash drive using the hash value (H) calculated in Step 10 ($ENDIF = E_K(DIF)$).

Step 12. The management program continues Step 11 with the whole DIF (ENDIF).

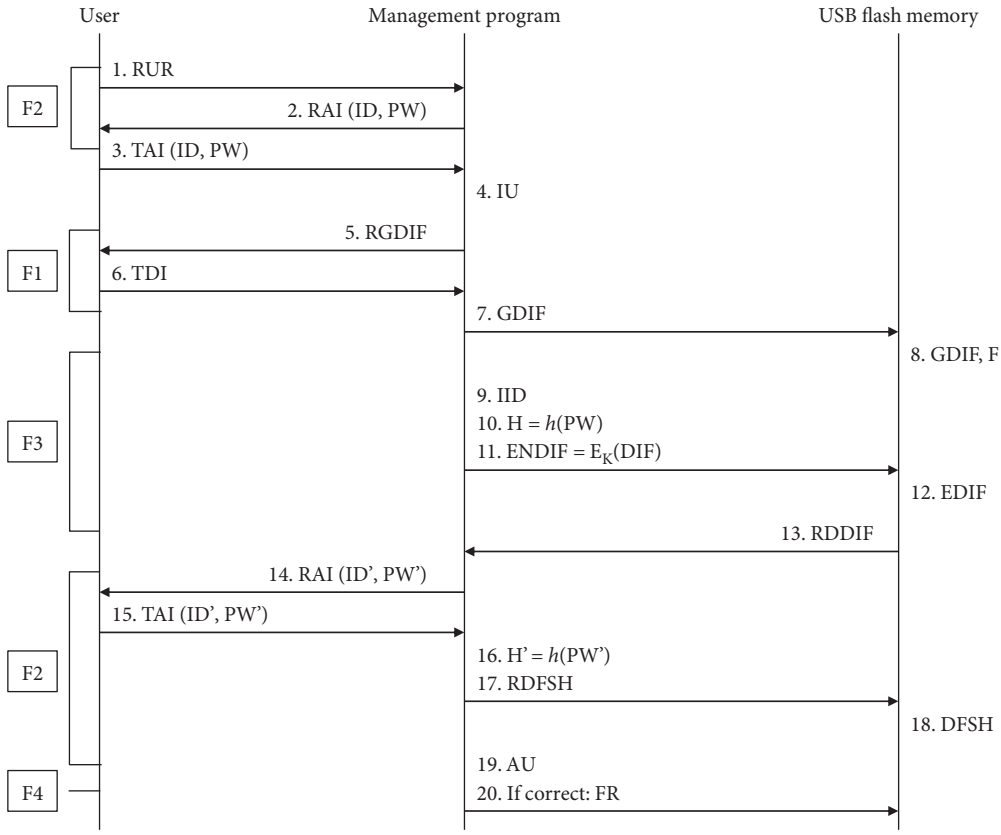


Figure 5: Proposed mechanism registration process.

Step 13. The first registration process is completed after Step 12. However, the proposed mechanism performs the user authentication process to validate the registration process. The management program requests decryption of the disk to the flash drive for user authentication (RDDIF).

Step 14. The management program requests the authentication data ID' and PW' from the user since the user ID and PW are needed to decrypt the disk (RAI).

Step 15. The user sends the ID' and PW' that he/she has registered (TAI).

Step 16. In the registration process, the user is authenticated by comparing the ID' input by the user in Step 15 with the ID inserted into the file system header in Step 9. To extract the inserted ID, the ID inserted in the file system header of the encrypted disk file must be decrypted and the hash operation is performed using the received PW' to generate the decryption key ($H' = h(PW')$)

Step 17. The management program requests a decryption of the file system header using the hash value (H') calculated in Step 16 (RDFSH).

Step 18. The decryption of the encrypted file system header of the disk image file is performed using the H' calculated with the user-input PW' as the key (DFSH).

Step 19. After the decryption of the encrypted file system header, the ID in the decrypted header and the ID' received from the user are compared for user authentication. The user is authenticated as a normal user if they match (AU) and the registration process is completed (FR). On the other hand, it is judged that an invalid password has been input if they do not match and the registration is canceled. The registration can be canceled by failed user authentication not only because of a malicious attacker attempting authentication bypass but also because of an error by a normal user. In that case, canceling registration with just an error will be very inefficient since the user will have to repeat the whole process starting from Step 1. To supplement this inefficiency, the user can return to Step 14 and input the authentication data again instead of repeating the whole process from Step 1. However, the process will be vulnerable to a brute force attack if infinite reentries of the authentication data are allowed. The input count can be limited to improve security, and the proposed mechanism limits the input count to 5 since the general input count is 5. If the input count exceeds 5, the user is judged to be a malicious user, and thus, the registration is canceled and all data are deleted.

The objects that participate in the registration process of the proposed mechanism are the user, the management program, and the USB flash drive. The data possessed by an object are deduced after the registration process is completed as the object information exposed to a third party can be used for malicious purposes. Although a user possesses the ID and

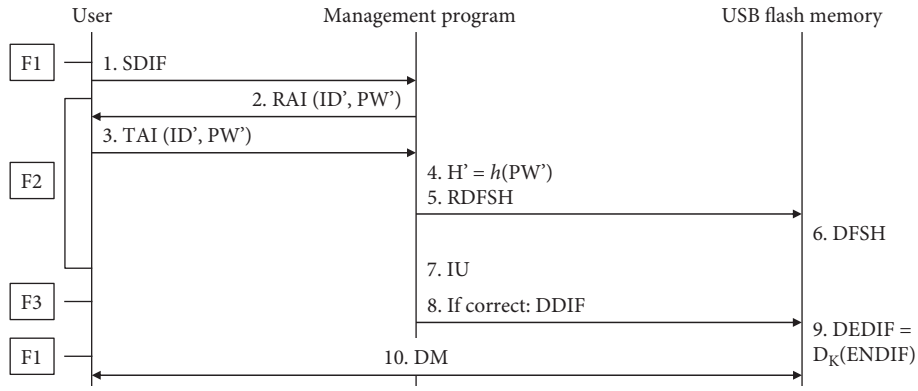


Figure 6: Proposed mechanism authentication process.

PW, which are the authentication data needed for registration, these data can be remembered only by the user and thus cannot be misused by attackers. Since the management program deletes the authentication data and the encryption/decryption data used in the registration process, there are no possessed data that can be misused by an attacker. Although the USB flash drive possesses the DIF encrypted with the hash value of the PW input by the user, the correct PW is remembered only by the user and the hash value of the correct PW cannot be inferred reversely because of the one-way characteristics of hash operation. Therefore, an attacker cannot abuse it. The applications using cryptology are all vulnerable to a brute force attack, but this problem can be partially resolved by limiting the input count and increasing the cryptologic strength.

When a user wants to use the disk after completing the above registration process, the following authentication process is performed to decrypt the encrypted DIF based on user authentication.

3.2. Authentication Process. The authentication process of the proposed mechanism allows only authorized users to use the encrypted disk. In more detail, user authentication is the process of authenticating the user by comparing the ID, which is the authentication data of the user requesting authentication, with the ID inserted into the file system header in the registration process. The file system header encrypted during the registration process is decrypted using the hash value of the PW received during the authentication process, and the user is authenticated when the data match. Then the whole disk file is decrypted and can be recognized as a general disk. Figure 6 shows the proposed authentication process.

Step 1. The user selects an EDIF to be decrypted from among the encrypted image files in the USB flash drive (SDIF).

Step 2. The management program requests the authentication data ID' and PW' to the user in order to decrypt the EDIF (RAI).

Step 3. The user sends the requested authentication data ID' and PW' to the management program (TAI).

Step 4. The management program performs the hash operation of the PW' input by the user in Step 3 to decrypt the encrypted file system header (H' = h(PW')).

Step 5. The management program requests the decryption of the file system header using the hash value (H') calculated in Step 4 (RDFSH).

Step 6. The decryption of the encrypted file system header of the disk image file is performed using the H' calculated with the user-input PW' as the key (DFSH).

Step 7. After the file system header has been decrypted, the ID inserted into the file system header in Step 7 is compared with the ID' input by the user in Step 3 to authenticate the user (IU). The user is authenticated and the whole DIF is decrypted (DDIF) if the comparison matches. Otherwise, it is judged that the IU has failed and the decryption of the DIF is canceled.

Step 8. After the decryption of the DIF, the file is mounted as a disk and can be used as a conventional disk after completing the mounting.

The above steps complete the whole authentication process. When the user no longer uses the disk, the decrypted flash drive is encrypted again so as to be protected safely. However, reusing the hash value generated in Step 4 can pose vulnerability since the key can be extorted by an attack, like memory scan, because the hash value of the PW, which is the encryption/decryption key, is continuously loaded in the memory while the disk is being used. Therefore, the encryption/decryption key allocated to the memory is deleted once the above authentication process has been completed. If the user does not use the disk afterward, the authentication data are requested again and used to encrypt the decrypted flash drive to protect the data when the drive is lost.

As with the registration process, the objects that participate in the authentication process of the proposed mechanism are the user, the management program, and the USB flash drive. The data possessed by an object are deduced since the object information exposed to a third party can be used

Table 7: Functions of the proposed mechanism and related steps.

Function	Detailed function	Related steps
F1 (disk management)	(i) File generation (ii) File opening (iii) File closing (iv) File deletion	(i) Registration process: Steps 5, 6, 7, and 8 (ii) Authentication process: Steps 1 and 8
F2 (user authentication)	(i) ID/PW (ii) Possible application of additional device authentication to improve security	(i) Registration process: Steps 1, 2, 3, 4, 14, 15, 16, 17, 18, and 19 (ii) Authentication process: Steps 2, 3, 4, 5, 6, and 7
F3 (data protection)	(i) Data encryption/decryption (ii) Saving of additional random number and use of hash chain to improve security	(i) Registration process: Steps 9, 10, 11, 12, 13, and 19 (ii) Authentication process: Step 7

for malicious purposes. Although a user possesses the ID and PW, which are the data needed for authentication, these data can be remembered only by the user and thus cannot be misused by attackers. Since the management program deletes the ID and password input by the user for decryption of the disk and comparison of the authentication data and the hash value of the PW, rather than storing them inside the management program after the completion of the authentication procedure, there are no data to be possessed, and thus, none can be used by an attacker. The USB flash drive does not save the user's PW and the encryption/decryption key during and after authentication; thus, an attacker cannot use the information.

The proposed mechanism provides disk management, user authentication, and data protection functions that use the registration process and the authentication process. It satisfies the requirements for data encryption/decryption, user authentication and identification, the prevention of arbitrary data copying, and the erasure of data for protection after loss of a drive, which are the four essential functions required to secure a USB as defined by the National Intelligence Service. Table 7 shows the steps related to these functions.

The data encryption/decryption requirement is satisfied with data protection function F3. To overcome the vulnerability of data being easily extorted because the original data are saved, as is the case in the existing USB flash drive, the image file to be used as a disk is generated and the generated DIF is encrypted and decrypted using the hash value of the user-input PW to prevent data leakage to the outside. The user authentication and identification requirement is satisfied with user authentication function F2. To overcome the vulnerability of data being extorted by authentication data exposure and authentication bypass of existing USB flash drives, it applies a measure to prevent the exposure of authentication data to prevent data leakage by authentication data exposure and authentication bypass. The prevention of arbitrary data copying requirement is satisfied by disk management function F1. After the image to be used as a disk is generated, it is mounted for use like a general disk only once the user has been authenticated. A user who fails the authentication procedure cannot use the disk since the disk is not mounted. The erasure of data for protection after loss is satisfied with user authentication function F2 and data protection function F3. Since all data are erased if an incorrect password is input more than 5 times, the leakage of data stored inside the flash drive is prevented even after the loss of a drive.

3.3. Security Evaluation. This section analyzes the security of the mechanism proposed in this paper. The analysis shows that the mechanism satisfies the confidentiality, integrity, authentication, and access control required by secure USB flash drives and safely protects the data from impersonation, man-in-the-middle, resending, and eavesdropping attacks by malicious attackers. The formalized verification tool AVISPA (Automated Validation of Internet Security Protocols and Applications) was used for the security analysis and the scenario based on the requirements and attack technologies described above.

3.3.1. Confidentiality. Secure USB flash drives require confidentiality in order to provide the internally stored data only to permitted users. Since the proposed mechanism uses the hash value of the user password to encrypt the whole file to be used as a disk, an attacker who does not have the password information cannot normally access the internally stored data.

3.3.2. Integrity. Secure USB flash drives must be able to guarantee integrity so that an unauthorized user cannot alter the internally stored data. Since the proposed mechanism uses the hash value of the user password to encrypt the whole file to be used as a disk, an attacker who does not have the password information cannot normally decrypt the data, and thus, integrity is assured.

3.3.3. Authentication. Secure USB flash drives must be able to authenticate the users to prevent unauthorized users from accessing the internally stored data and from normally accessing the internally stored data even if authentication is bypassed. The proposed mechanism authenticates users based on the ID and PW and encrypts the ID using the hash value of the PW. Therefore, the data are decrypted into unrecognizable data if an attacker inputs an arbitrary ID and PW to bypass the authentication. Since the disk is decrypted with the hash value of the arbitrary PW if an attacker bypasses authentication with the arbitrary data, the resulting disk data will be invalid, and thus, the internally stored original data will not be accessed.

3.3.4. Access Control. Secure USB flash drives must be able to prevent unauthorized users from accessing the internally stored data. Since the proposed mechanism encrypts the whole disk using the hash value of the user PW, an attacker who does not have the correct password or the hash value of the PW cannot obtain the correct decryption key and thus cannot decrypt the disk. Since the attacker cannot decrypt the disk, it is impossible to access the original data stored inside.

3.3.5. Impersonation Attack. Secure USB flash drives must be able to prevent an attacker from impersonating an authorized user so as to recover the internally stored data. Although an attacker may obtain the encrypted file header and the whole encrypted disk, an attacker who does not have the user PW cannot normally decrypt the disk since the decryption key is based on the user PW. Therefore, the proposed mechanism can protect the internally stored data from impersonation attacks.

3.3.6. Man-in-the-Middle Attack. Secure USB flash drives must be able to prevent man-in-the-middle attacks from extorting the authentication data and encryption/decryption key by inserting an additional module in the front part. Although an attacker may obtain the encrypted file header, encrypted ID, and encrypted disk file, the attacker will not have the information needed to decrypt the encrypted data since the user PW and the hash value of the PW used as the key are not transferred to the flash drive but are utilized only inside the management program. Therefore, the proposed mechanism is safe from man-in-the-middle attacks.

3.3.7. Eavesdropping Attack. Secure USB flash drives must be able to prevent eavesdropping attacks from obtaining and abusing the data transferred between the host and the flash drive. Since the proposed mechanism encrypts all data transferred between the host and the flash drive, it does not expose the authentication data and the encryption/decryption key and thus is safe from eavesdropping attacks.

3.3.8. Resending Attack. Secure USB flash drives must be able to prevent resending attacks which obtain the data transferred between the host and the flash drive and then resends them to bypass the authentication and access the data. Since the proposed mechanism only transfers the encrypted disk data from the flash drive, an attacker cannot bypass authentication and access the data even after obtaining the information. Therefore, the proposed mechanism is safe from resending attacks.

Lastly, the formalized verification tool AVISPA was used for the security analysis; the results of which showed that the proposed mechanism is safe. Figures 7, 8, 9, and 10 show the code used for the analysis and the analysis results.

```
protocol SecureUSB;
Identifiers
A, B                : user;
ID, FSH, DIF        : number;
Kps                 : symmetric_key; %Hashed PW

messages
1. A → B            : {ID, FSH}Kps
2. A → B            : {DIF}Kps
3. B → A            : {ID, FSH}Kps
4. B → A            : {DIF}Kps

knowledge
A                   : A, B, Kps:
B                   : A, B, Kps

session_instances
[A: program,B:usb,Kps:hashedpw];

goal
secrecy_of ID[];
secrecy_of DIF[];
A authenticates B on FSH;
```

Figure 7: AVISPA CAS code.

```
SUMMARY
  SAFE

DETAILS
  BOUNDED_NUMBER_OF_SESSIONS
  TYPED_MODEL

PROTOCOL
  C:\progra~1\SPAN\testsuite\results\hlpslGenFile.if

GOAL
  As Specified

BACKEND
  CL-AtSe

STATISTICS
  Analysed: 5 states
  Reachable: 3 states
  Translation: 0.00 seconds
  Computation: 0.00 seconds
```

Figure 8: AVISPA ATSE result.

4. Conclusion

The USB flash drive is currently the most popular mobile storage unit because of its many strengths including fast data transfer speed, high portability, and free transfer and deletion. However, serious problems have arisen, such as the inability to protect the internally stored data after the loss of a USB drive, leading to demands for the development of a secure USB flash drive featuring improved security functions. For that reason, new and more secure USB flash drives protect the internally stored data using such security technologies as data encryption/decryption and user authentication

```
%OFMC
% Version of 2006/02/13
SUMMARY
  SAFE
DETAILS
  BOUNDED_NUMBER_OF_SESSIONS
PROTOCOL
  C:\progra~1\SPAN\testsuite\results\hlpslGenFile.if
GOAL
  as_specified
BACKEND
  OFMC
COMMENTS
STATISTICS
  parseTime: 0.00s
  searchTime: 0.01s
  visitedNodes: 10 nodes
  depth: 5 plies
```

Figure 9: AVISPA OFMC result.

```
SUMMARY
 INCONCLUSIVE
DETAILS
 NOT_SUPPORTED
PROTOCOL
 C:\progra~1\SPAN\testsuite\results\hlpslGenFile.if
GOAL
 SECRECY
BACKEND
 TA4SP
COMMENTS
 Some rules may be not fired so TA4SP does not do the verification.
STATICTICS
 Translation: 0.00 seconds
```

Figure 10: AVISPA TA4SP result.

and identification. However, the problems of access to the inside of a drive and the leakage of data have been identified in secure USB flash drives installed with the latest security technologies due to such vulnerabilities as implementation and environmental vulnerabilities, unlock command, and reverse engineering. To solve such problems, this paper proposes a safe secure USB flash drive mechanism that does not expose the authentication data. The mechanism overcomes the existing vulnerabilities to protect the data more safely, since it does not store the data needed for user authentication and disk decryption inside the flash drive data and has no routine for comparing the authentication. To analyze the security of the proposed mechanism, the security requirements which the secure USB flash drive must satisfy and an attack technology scenario were deduced. The results of the security assessment confirmed that the proposed mechanism satisfies the confidentiality, integrity, authentication, and access control requirements and safely protects the data from impersonation, man-in-the-middle, resending, and eavesdropping attacks. In addition, the formalized verification tool AVISPA was used for the security analysis; the results of which showed that the proposed mechanism is indeed safe.

Although the mechanism proposed in this paper was applied to a secure USB flash drive, it could be used in other areas. For example, it could be applied to a secure disk to improve the security of a hard disk and to secure backup storage so as to safely back up data and thereby protect internally stored data. Planned future studies include the application of a more improved mechanism, such as by adding device authentication to improve security through access control of the disk itself and by using the hash chain and a random number to supplement the weakness entailed by encryption with the same key.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Disclosure

A part of this paper was presented at a conference on the International Symposium on Mobile Internet Security (MobiSec), October 19–22, 2017, Jeju Island, South Korea.

Conflicts of Interest

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

Acknowledgments

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) that is funded by the Ministry of Education (NRF-2015R1D1A1A01057300) and by the Soonchunhyang University fund.

References

[1] Wikipedia, "USB Flash Drive," August 2016, https://ko.wikipedia.org/wiki/USB_%ED%94%8C%EB%9E%98%EC%8B%9C_%EB%93%9C%EB%9D%BC%EC%9D%B4%EB%B8%8C.

[2] K. Lee, K. Yim, and E. H. Spafford, "Reverse-safe authentication protocol for secure USB memories," *Security and Communication Networks*, vol. 5, no. 8, p. 845, 2012.

[3] O. Insu, Y. Lee, H. Lee, K. Lee, and K. Yim, "Study on secure USB mechanism without exposure of the authentication information," in *Proceedings of the International Symposium on Mobile Internet Security (MobiSec)*, Jeju Island, South Korea, October 2017.

[4] K. Lee, H. Yeuk, Y. Choi, S. Pho, I. You, and K. Yim, "Safe authentication protocol for secure USB memories," *Journal of Wireless Mobile Networks, Ubiquitous Computing and Dependable Applications*, vol. 1, no. 1, pp. 46–55, 2010.

[5] J. Bang, B. Yoo, and S. Lee, "Secure USB bypassing tool," *Digital Investigation*, vol. 7, pp. S114–S120, 2010.

[6] S.-H. Lee, K.-B. Yim, and I.-Y. Lee, "A secure solution for USB flash drives using FAT file system structure," in *2010 13th International Conference on Network-Based Information Systems*, pp. 487–492, Takayama, Japan, September 2010.

[7] H. Jeong, Y. Choi, W. Jeon et al., "Vulnerability analysis of secure USB flash drives," in *2007 IEEE International Workshop on Memory Technology, Design and Testing*, pp. 61–64, Taipei, Taiwan, December 2007.

[8] J. Kim, Y. Lee, K. Lee, T. Jung, D. Volokhov, and K. Yim, "Vulnerability to flash controller for secure USB drives," *Journal of Internet Services and Information Security*, vol. 3, no. 3/4, pp. 136–145, 2013.

[9] K.-G. Lee, H.-W. Lee, C.-W. Park, J.-W. Bang, K.-y. Kim, and S. Lee, "USB PassOn: secure USB thumb drive forensic toolkit," in *2008 Second International Conference on Future Generation Communication and Networking*, pp. 279–282, Hainan Island, China, December 2008.

[10] S.-H. Lee, J. Kwak, and I.-Y. Lee, "The study on the security solutions of USB memory," in *Proceedings of the 4th International Conference on Ubiquitous Information Technologies & Applications*, pp. 1–4, Fukuoka, Japan, December 2009.

[11] A. N. Magdum and Y. M. Patil, "A secure data transfer algorithm for USB mass storage devices to protect documents," *International Journal of Emerging Engineering Research and Technology*, vol. 2, no. 4, pp. 113–119, 2014.

[12] L. Hamid, "Biometric technology: not a password replacement, but a complement," *Biometric Technology Today*, vol. 2015, no. 6, pp. 7–10, 2015.

[13] S.-H. Lee and I.-Y. Lee, "Secure index management scheme on cloud storage environment," *International Journal of Security and Its Applications*, vol. 6, no. 3, pp. 75–82, 2012.

Hindawi
Journal of Sensors
Volume 2018, Article ID 1386470, 7 pages
https://doi.org/10.1155/2018/1386470

Research Article

Fog Computing-Based IoT for Health Monitoring System

Anand Paul[1], **Hameed Pinjari**[1], **Won-Hwa Hong**[2], **Hyun Cheol Seo**[2], **and Seungmin Rho**[3]

[1]*School of Computer Science and Engineering, Kyungpook National University, Daegu, Republic of Korea*
[2]*School of Architectural, Civil, Environmental and Energy Engineering, Kyungpook National University, Daegu, Republic of Korea*
[3]*Department of Media Software, Sungkyul University, Anyang, Republic of Korea*

Correspondence should be addressed to Anand Paul; paul.editor@gmail.com and Won-Hwa Hong; hongwonhwa@gmail.com

Received 30 October 2017; Accepted 3 May 2018; Published 22 October 2018

Academic Editor: Carmine Granata

Wireless sensor networks (WSNs) are widely used in the area of health informatics. Wireless and wearable sensors have become prevalent devices to monitor patients at risk for chronic diseases. This helps ascertain that patients comply by the treatment plans and also safeguard them during sudden attacks. The amount of data that are gathered from various sensors is numerous. In this paper, we propose to use fog computing to help monitor patients suffering from chronic diseases such that the data are collected and processed in an efficient manner. The main challenge would be to only sort out context-sensitive data that are relevant to the health of the patient. Just having a simple sensor-to-cloud architecture is not viable, and this is where having a fog computing layer makes a difference. This increases the efficiency of the entire system, as it not only reduces the amount of data that is transported back and forth between the cloud and the sensors but also eliminates the risk that a data center failure bears with it. We also analyze the security and deployment issues of this fog computing layer.

1. Introduction

With each passing day, the way the world interacts changes. The past few years have seen an increase in the usage of cloud computing for a huge number of applications. But now, the cloud does not cater to just simple applications and technological needs; the Internet of things (IoT) is proving to be the next technological trend when it comes to cloud computing. In healthcare applications, wireless sensor networks (WSN) have started playing a huge role in the way patients are being monitored. Wireless sensors in the form of wireless wearable accessories or devices are attached to a patient such that this information can be used for the monitoring process. The sensors can be of various forms and sizes [1, 2] as long as they are relevant to the need.

The wireless sensor networks generate a huge amount of data. These data that have been collected from all the devices connected to the network may be useful as well as redundant. All these unprecedented amounts of data can overwhelm the data storage systems and the data analysis applications. The weeding out of irrelevant data has to be a context-sensitive process. Hence, the sensors would have to send the data collected to computing devices that are capable of performing tasks of analysis, aggregation, and storage. In many cases, each patient requires a high number of sensors, and hence, creating an infrastructure dedicated to an individual becomes inefficient. Hence, IoT provides an alternative approach in which sensor devices are used in a common infrastructure. These sensor devices can then forward the data to a cloud server.

For many healthcare applications, having a simple sensor-to-cloud architecture is not viable, especially due to the fact that most hospitals would not prefer patient data to be stored outside [3]. Also, there is always the bleak case of there being a network failure or a data center failure, which puts patients' health at risk. This is where fog computing aids healthcare applications.

Using only the cloud may cause delay during the transfer of data from the sensors to the cloud and the cloud to the hospitals or personal physicians. In healthcare, we have

emergency response systems that require real-time operations in which efficiency and time play an important part; this may suffer due to delay caused by the cloud [4]. Hence, transfer of such immense amounts of data back and forth is not an efficient option not only due to latency issues but also due to security. The risks involved here are not only infringement of data but also risks to the health of patients. Hence, the classical centralized cloud computing architecture has to be extended to a distributed one. A distributed architecture refers to tasks being divided and then offloaded to more than one node. This is commonly referred to as edge computing.

In edge computing, the main computationally intensive operations are performed at the edge of the network instead of holding it in the cloud or on a centralized data warehouse. Edge computing uses computing resources near IoT devices for local storage and preliminary data processing. According to Cisco [5], by 2020, 50 billion devices will be connected to the Internet. Hence, edge computing will also require greater flexibility in order to manage this huge influx of devices. Edge devices cannot handle multiple IoT applications competing for their limited resources, which results in resource contention and increases processing latency. A distributed feature added to this will help in scalability and reduces the risk of exposure of data and hence increases security and eliminates most of the privacy concerns [6].

A fog computing layer integrates edge devices with cloud resources and hence extends the existing cloud infrastructure [7]. In this architecture, the application resides not only in the cloud but also in the devices closest to the patients and the infrastructure components between them. The term infrastructure refers to access points, gateways, and routers. The main objective of healthcare applications is to provide constant supervision on the health of a patient. Implementing this fog layer provides for the successful fulfillment of this requirement. Also, due to the fact that data are stored in data centers in the cloud, the problem of infrastructure, maintenance, upgrades, and costs is solved.

In this paper, we strive to improve current healthcare systems by implementing a context-sensitive fog computing environment. We discuss the various computational tasks involved in healthcare that will be performed in the fog layer, the cloud, or at the user devices and sensors. The security of the information passing the system will be improved as the exposure of data is limited due to the fact that it does not have to travel to and fro in the network. The services that the cloud performs can be distributed to other nodes in such a way that the overhead time taken for data to go back and forth from the cloud to devices is compensated or reduced. Distributing the services among various nodes overcomes some of the challenges that previous healthcare monitoring applications have faced.

The remainder of the paper is organized as follows. In Section 2, we discuss about fog computing, the expectations, and various challenges that have to be faced in the process of implementing fog computing in healthcare. The proposed research is discussed in Section 3. Section 4 analyzes the proposed scheme and the experimental results obtained. The conclusion to the work is given in Section 5.

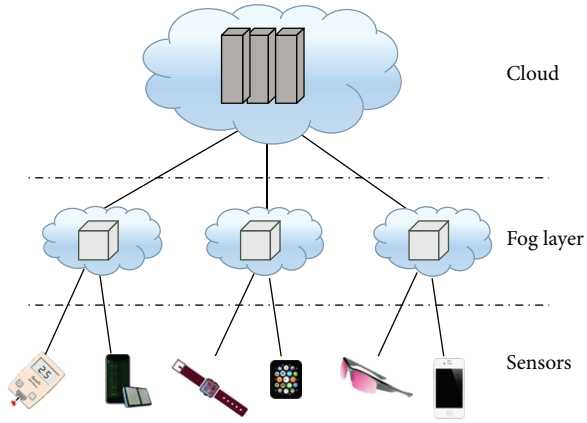


Figure 1: The layers in IoT.

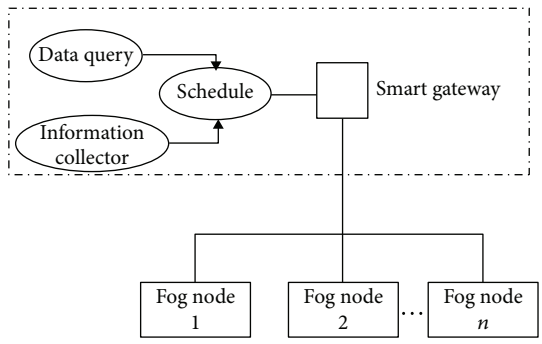


Figure 2: The architecture of a fog layer.

2. Related Work

In this section, we discuss the various technologies that will be used in the proposed research. Also, the research already done in this area will be reviewed in the upcoming subsections.

2.1. Fog Computing. Fog computing is highly virtualized and provides a medium for computing, storage, and networking between end devices and the cloud [7]. The main notion of fog computing is to migrate the tasks of data centers to fog nodes situated at the edge of the network. We refer to these fog nodes as the fog layer. As these devices that perform the tasks are at the edge of the network, it results in a higher data transfer rate and a reduced user response time.

Figures 1 and 2 show the architecture of a simple fog layer with respect to the cloud and the sensors.

2.2. Research in Fog Computing. The main applications where fog computing will prove very useful are time-sensitive applications where a huge amount of data has to be processed [8]. The following are the application areas that can benefit from fog computing.

(1) Healthcare applications: this is an area in which real-time processing plays an important role. Hence, data will have to be processed very fast, and the response time will have to be as less as possible [4, 9, 10].

(2) Augmented reality: for augmented reality (AR), overlaying useful information onto the physical world in real time is very important. Using fog computing will help achieve this aspect of augmented reality [11].

(3) Smart utility services: utility services like electricity, water, telephone, and so on can be managed by fog computing.

(4) Traffic management system: fog computing can increase the efficiency of the traffic signal system by reducing the latency. The interactions between vehicles, traffic signals, and access points can be enhanced by the fog [12].

(5) Caching and processing: fog computing can also be used for improving the performance of websites [13]. Certain websites have a lot of databases and data to be processed, for example, social networking sites and library or online shopping malls. These websites can use the fog layer for caching and preprocessing its data and hence reducing time and space complexity.

(6) Gaming: in the past few years, there has been a huge evolution in the gaming industry. Apart from games being computationally complex, they are mostly multiplayer these days and depend greatly on real-time processing.

(7) Decentralized smart building control: similar to smart utility services, even in the case of smart building, control fog computing will play a huge role in making it more efficient and secure.

Fog computing offers enormous advantages for delay-sensitive fog-based application. Chen et al. [14] implemented a prototype of a smart gateway for the use of WSN in healthcare systems at home. The system is able to transmit reports at real time in a low power embedded system. Hong et al. [15] presented Mobile Fog, which is a programming model for Internet applications that are geographically distributed and latency-sensitive.

2.3. Motivation. The use of smart devices has fortuitously been exercised in healthcare. These days, it is commonplace to find a range of healthcare gadgets that can be used by patients at home or even worn by them. The gadgets mostly encompass sensors. These sensors generally incorporate transducers and are capable of detecting electrical, thermal optical, chemical, and other signals [10].

The main motivation of this paper is to enhance health monitoring systems that are based on IoT devices such that the information collected from WSNs are processed efficiently, and the context-sensitive data that are relevant to the patient are considered. For this, we implement a fog layer that improves the latency of health monitoring systems and enables real-time health monitoring. By this, we strive to also ensure security for the information of the patients such that patient privacy is maintained, and also, tampering of data by third party is avoided.

2.4. Key Objectives of Healthcare Systems. Some of the key objectives of healthcare systems are [16]

(1) improved clinical decision making,

(2) reduced duplication of diagnostic testing, imaging, and history taking,

(3) better medication management,

(4) increased adoption of screening programs and preventive health measures.

2.5. Challenges in Implementing a Health Monitoring System. In order to successfully implement a health monitoring system, a number of challenges have to be addressed [16]:

(1) Safeguarding privacy and security

(2) Technical problems

(3) Organizational barriers

(4) Financial costs

(5) Different policies

(6) Training programs for practitioners and healthcare providers

3. Proposed Research

In this section, we present a tri-tier architecture for context- and latency-sensitive health monitoring using cloud and fog computing. The tri-tiers consist of cloud computing, fog computing, and sensors which work in conjunction with one another. Sensors consist of wearable or not wearable devices that are attached to the patients in the form of smart watches, fitness bands, smart phones, wearable glasses, and so on. The applications used for health monitoring will have components running in the edge devices situated in the fog layer, the wearable sensors, or the cloud. The edge devices may be controlled by the cloud and the fog layers. Information will flow across this tri-tiered infrastructure.

In context-sensitive health monitoring, personalized care can be given to each patient. Context can be classified into extrinsic and intrinsic context. In case of healthcare, extrinsic context is influenced by external factors like the environment that is surrounding the patient. Environmental sensors can be used to extract the extrinsic parameters of a user, and the intrinsic context can be extracted by biosensors. Both intrinsic and extrinsic sensors provide relevant information that may be used for monitoring patients' health. But depending on the disease of the patient, the type of data that is relevant differs. Hence, maintaining context-sensitive data processing has to also be done by the fog layer.

The schematic diagram of the proposed architecture is given in Figure 3. In this architecture of context-and latency-sensitive health monitoring systems, we ensure that all the key objectives of health monitoring and more are all accomplished.

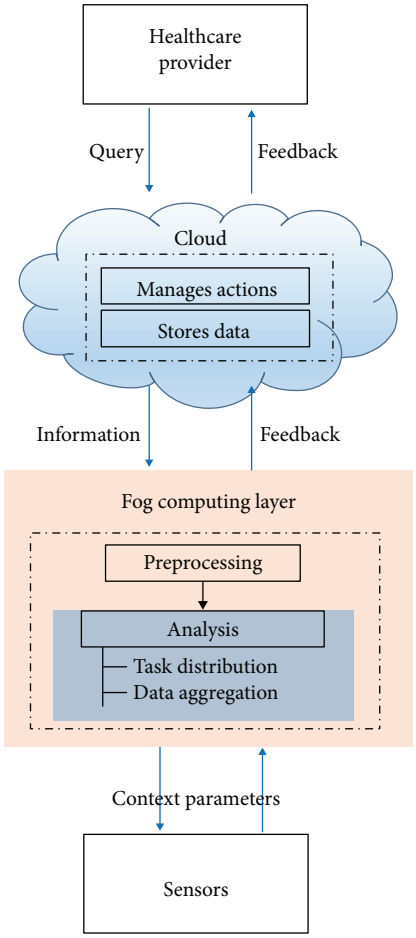


Figure 3: Schematic diagram of proposed architecture.

3.1. Sensors Tier. These are the devices that gather information from the patients. These sensors gather both extrinsic and intrinsic values. Extrinsic characteristics are the temperature, location, and so on. Intrinsic characteristics are the blood pressure, blood glucose level, heartbeat, and so on that are collected by the patient's wearable sensors. The patient can also enter data into his or her smart phone, and these data will then be made available for processing. The job of the sensors is to collect all this data and send them to the fog computing layer.

3.2. Fog Computing Tier. This layer performs data analysis and aggregation of the data. The data and information collected by the edge devices are analyzed in this layer. This layer behaves as the server. Massive amounts of real-time data from sensors are sent to this tier. The fog layer then distributes the processing work to various edge devices connected to the fog layer, and hence, massive amounts of data are analyzed. The distribution of the processing work has to be done using an efficient task-scheduling algorithm.

(1) Work distribution: this task is performed via the smart gateway using a scheduler. In this paper, we distribute the tasks using the task-scheduling mechanism for cloud-fog computing systems introduced by Pham and Huh [17]. Two graphs are first created: task graph and processor graph. Figure 4 gives a graphical representation of the task and processor graphs.

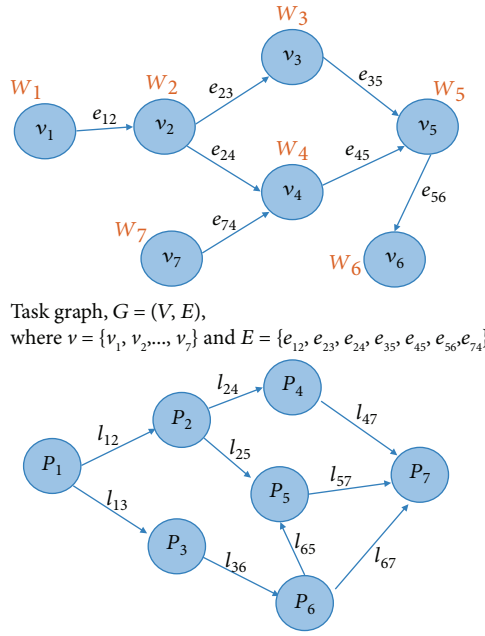


Figure 4: Graphical representation of the task and processor graphs.

Let $G = (V, E)$ represent the task graph, where G is a directed acyclic graph (DAG); V is the set of vertices $\{v_1, v_2, \ldots, v_n\}$ that denote the parallel subtasks, and E is the set of edges where each edge $e_{ij} \in E$ implies that task v_i has a corresponding workload w_i that signifies the amount of work to be processed at a particular resource. Every edge e_{ij} has a corresponding weight c_{ij} which represents the amount of data that is transferred from v_i to $v + i$.

Let $H = (R, L)$ be a DAG that represents the processor graph, where R denotes the set of vertices $\{P_1, P_2, \ldots, P_n\}$ where each $P_i \in R$ is a processor at the cloud or fog. The edge $l_{ij} \in L$ denotes a link between processor P_i and P_j. Now, $R = N_{\text{cloud}} \cup N_{\text{fog}}$, where N_{cloud} and N_{fog} denote the set of cloud nodes and fog nodes, respectively.

The priority of computing a particular task is calculated as

$$\text{pri}(v_i) = \begin{cases} \overline{w(v_i)} & \text{if } v_i \neq v_{\text{exit}}, \\ \overline{w(v_i)} + \max_{v_j \in \text{succ}(v_i)} \left[\overline{c(e_{ij})} + \text{pri}(v_j) \right], & \text{if } v_i = v_{\text{exit}}, \end{cases} \quad (1)$$

where $\overline{w(v_i)} = w(v_i)/\overline{w(v_i)}$, and $\overline{c(e_{ij})} = c_{ij}/\overline{BW}$; also, BW is the bandwidth.

Table 1: Comparison of percentages.

Physical topology	Average latency (ms)		Network usage (KBs)	
	Cloud only	*With fog layer*	*Cloud only*	*With fog layer*
Config 1	210.38	8.47	130	12
Config 2	210.78	8.47	351	22
Config 3	211.57	8.47	672	53
Config 4	1283.86	8.47	1061	98
Config 5	3225.91	8.47	1102	189

Once the priority for the various tasks has been calculated, these tasks are sent to nodes for execution. Now the choice of the nodes to do a particular task has to be figured out. For this, the time required by the processor at each node and the processing speed have to be taken into consideration. The earliest start time and earliest finish time will be used for making those calculations. For more details, please refer to the paper by Pham and Huh [17].

(2) Data aggregation: once tasks are distributed, the data have to be aggregated. Data aggregation consists of three main parts: schema mapping, duplicated detection, and data fusion. Schema mapping will ensure that the data are aggregated in such a way that it makes sense, and there is a flow to the data. Duplicated detection ensures that there will not be any redundancy in the data. Redundancy is a challenge because there are many nodes that perform the work, as a result, there may be overlap in data. Apart from duplicate detection, false data injection is also avoided to ensure security in the fog device. This is implemented by adding a local filter to the fog device. Data fusion is the final stage of data aggregation in which the final information is gathered and put together as one entity.

3.3. Cloud Computing Tier. This is the layer that manages the various actions that are to be performed by the health monitoring system. A component of the monitoring app runs on the sensors which enables the sensors to collect data and send it to the fog layer. The decision of what task to perform is done with the help of the fog layer as described in the previous subsection. The cloud computing tier constantly supervises the health monitoring system.

3.4. Health Monitoring System. Apart from the various tiers described above, the healthcare delivery system typically consists of four levels: the region, the institution, the clinical department or outpatient clinic, and the individual physician, nurse, or patient [18]. The flow of information between these four levels has to be efficiently managed. Hence, there are many important privacy and security challenges that have to be met.

4. Experimental Results and Analysis

In this section, we discuss the experiments performed so far and the analysis of the proposed system.

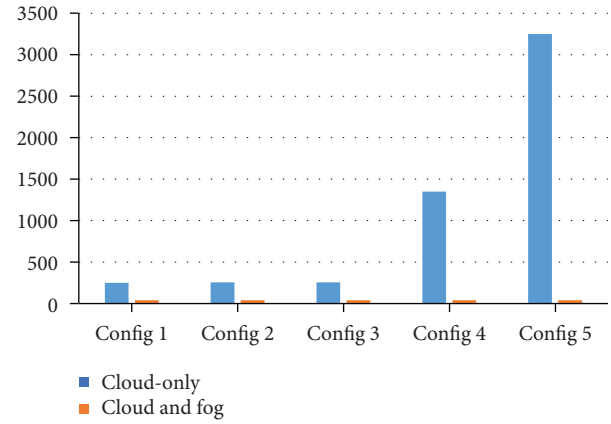


Figure 5: Average latency comparison.

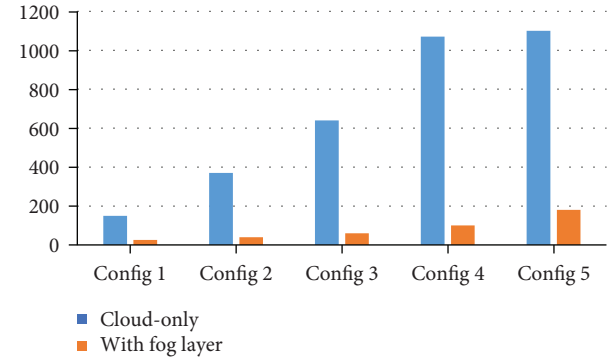


Figure 6: Network usage comparison.

4.1. Simulation of the Fog Environment. For the purpose of comparing the use of fog computing versus just using the traditional cloud computing in healthcare, we used iFogSim toolkit to simulate the fog network [19].

The iFogSim simulates the specified configuration and gives the simulated outputs. This makes it convenient to observe end results when all required technology is not available. The simulator itself adds a bit overhead time. For this simulation, we conducted several test runs for 5 configurations of monitoring devices.

The average latency and network usage for the 5 configurations are given in Table 1. The table shows that the configurations of connected devices do not really affect the latency of our architecture that uses fog computing. The network usage of the architecture with fog computing is much lower than the architecture that uses only fog computing. The iFogSim toolkit is used to simulate the 5 different configurations. For each of the 5 configurations, the monitoring devices are varied. In Config 1, Config 2, Config 3, Config 4, and Config 5, they each have 4, 8, 16, 32, and 64 monitoring devices, respectively. So each configuration will give different results when simulated. The monitoring devices used in the configurations have a CPU length of 1000 million instructions, a network length of 20,000 bytes, and an average interarrival time of 5 ms. Figures 5 and 6 give the comparison of the latency and network usage for the various configurations. The simulation of configurations that use only the

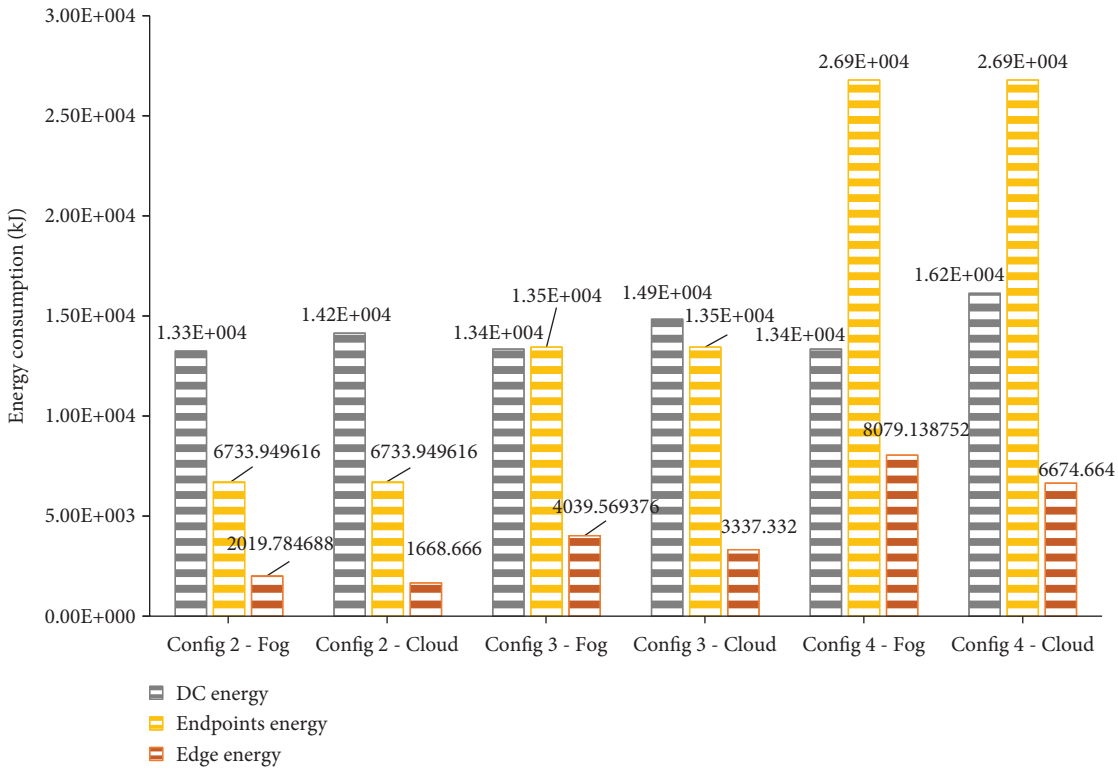


Figure 7: Comparison of energy comparison.

cloud is done using the CloudSim toolkit, and those that use the cloud in conjunction with the fog layer is done using the iFogSim toolkit. From both the charts, we can observe that the complexity of the fog layer does not affect the latency or the network usage. In fact, fog computing increases the efficiency of the entire system. From Table 1, we can observe that using fog layer indeed decreases the latency and network usage when compared to using only cloud computing. Figure 7 shows the energy consumption of fog computing versus only using the cloud. From the figure, it can be analyzed that the energy consumption for fog computing mainly takes places at the edge where most of the processing is done. On the other hand, in the case of cloud computing, the energy is mainly spent at the data centers or at the cloud.

4.2. Analysis

4.2.1. Latency. In our implementation of fog computing in health informatics, there will be transfer of data among the various tiers. The amount of data and the time taken will differ in different cases. Hence, the latency differs. Let us consider L_f as the latency when the evaluated data has to be returned to the IoT devices and L_e be the latency when the data is sent to the cloud. Then,

$$L_f = t_s + t_r + e_f,$$
$$L_e = t_s + e_f + e_e. \quad (2)$$

In the above equations, t_s is the time taken from IoT sensors to the fog layer; t_r is the time taken for the data to return from the fog layer to the IoT devices; e_f is the evaluation time taken by the edge devices; and e_e is the evaluation time taken at the cloud.

These two equations will be used by us to analyze the latency times. This is very important because most applications of fog computing rely on the real-time processing capabilities of the network.

4.2.2. Computation. The computations in the fog layer should be real-time and latency-sensitive services. Many techniques for reducing the computation complexities have to be adopted. The data packets can be stored at the fog nodes for some time to avoid reloading of the same data. According to some renewal algorithms, these data packets may be replenished by new data packets. Intelligently, distributing data packets to the most efficient number of edge devices also plays a very important role.

4.2.3. Security Analysis. Introducing a fog layer into the cloud computing infrastructure reduces the security risk when it comes to data of the patients not getting lost due to failure in a data center. But at the same time, the data is stored in the cloud. This increases the threat to the privacy of patient information. In this scheme, we propose to secure the data of the patient by encrypting the data of the patient using a secret key.

5. Conclusion

With increase in the use of the IoT, there is a huge demand for data to be processed in real time and efficiently. The

interest in implementing fog computing as a technology is growing rapidly as shown by the H2020 ICT Work Programme for 2016-17. In this paper, we have proposed a scheme for fog computing in health monitoring systems. The work shown in this paper is a work in progress. There are many factors that have hindered and slowed down our implementation process. These factors include not readily available software, complexity, and so on. In spite of all these obstacles, the research area is a very promising one and when implemented will prove to be a very useful technology.

The future work is to deploy this system to various edge devices and judge the behavior of the system.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the National Research Foundation Grant funded by the Korean Government (NRF-2017R1C1B5017464). This study was also supported by the National Research Foundation of Korea (NRF) funded by the Korean government (MSIP) (NRF-2016R1A2A1A05005459).

References

[1] D. Kang, Y. S. Kim, G. Ornelas, M. Sinha, K. Naidu, and T. Coleman, "Scalable microfabrication procedures for adhesive-integrated flexible and stretchable electronic sensors," *Sensors*, vol. 15, no. 9, pp. 23459–23476, 2015.

[2] N. M. Farandos, A. K. Yetisen, M. J. Monteiro, C. R. Lowe, and S. H. Yun, "Contact lens sensors in ocular diagnostics," *Advanced Healthcare Materials*, vol. 4, no. 6, pp. 792–810, 2015.

[3] F. A. Kraemer, A. E. Braten, N. Tamkittikhun, and D. Palma, "Fog computing in healthcare–a review and discussion," *IEEE Access*, vol. 5, pp. 9206–9222, 2017.

[4] Y. Cao, S. Chen, P. Hou, and D. Brown, "FAST: a fog computing assisted distributed analytics system to monitor fall for stroke mitigation," in *2015 IEEE International Conference on Networking, Architecture and Storage (NAS)*, pp. 2–11, Boston, MA, USA, August 2015.

[5] Cisco, "White paper: fog computing and the internet of things: extend the cloud to where the things are," Tech. Rep., Cisco, 2015.

[6] L. M. Vaquero and L. Rodero-Merino, "Finding your way in the fog: towards a comprehensive definition of fog computing," *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 5, pp. 27–32, 2014.

[7] F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, "Fog computing and its role in the internet of things," in *Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing*, pp. 13–16, Helsinki, Finland, August 2012.

[8] S. Yi, C. Li, and Q. Li, "A survey of fog computing: concepts, applications and issues," in *Proceedings of the 2015 Workshop on Mobile Big Data*, pp. 37–42, Hangzhou, China, June 2015.

[9] K. Ha, Z. Chen, W. Hu, W. Richter, P. Pillai, and M. Satyanarayanan, "Towards wearable cognitive assistance," in *Proceedings of the 12th Annual International Conference on Mobile Systems, Applications, and Services*, pp. 68–81, Bretton Woods, NH, USA, June 2014.

[10] V. Stantchev, A. Barnawi, S. Ghulam, J. Schubert, and G. Tamm, "Smart items, fog and cloud computing as enablers of servitization in healthcare," *Sensors & Transducers*, vol. 185, no. 2, pp. 121–128, 2015.

[11] J. K. Zao, T. T. Gan, C. K. You et al., "Augmented brain computer interaction based on fog computing and linked data," in *2014 International Conference on Intelligent Environments*, pp. 374–377, Shanghai, China, 2014.

[12] F. Bonomi, R. Milito, P. Natarajan, and J. Zhu, "Fog computing: a platform for internet of things and analytics," in *Big Data and Internet of Things: A Roadmap for Smart Environments*, pp. 169–186, Springer, 2014.

[13] J. Zhu, D. S. Chan, M. S. Prabhu, P. Natarajan, H. Hu, and F. Bonomi, "Improving web sites performance using edge servers in fog computing architecture," in *2013 IEEE Seventh International Symposium on Service-Oriented System Engineering*, pp. 320–323, Redwood City, USA, March 2013.

[14] Y. Chen, W. Shen, H. Huo, and Y. Xu, "A smart gateway for health care system using wireless sensor network," in *2010 Fourth International Conference on Sensor Technologies and Applications*, pp. 545–550, Venice, Italy, July 2010.

[15] K. Hong, D. Lillethun, U. Ramachandran, B. Ottenwälder, and B. Koldehofe, "Mobile fog: a programming model for large-scale applications on the internet of things," in *Proceedings of the Second ACM SIGCOMM Workshop on Mobile Cloud Computing*, pp. 15–20, Hong Kong, China, August 2013.

[16] Institute of Medicine (US) Committee on Data Standards for Patient Safety, *Key Capabilities of an Electronic Health Record System: Letter Report*, National Academies Press, 2003.

[17] X.-Q. Pham and E.-N. Huh, "Towards task scheduling in a cloud-fog computing system," in *2016 18th Asia-Pacific Network Operations and Management Symposium (APNOMS)*, pp. 1–4, Kanazawa, Japan, October 2016.

[18] T. Benson, "Why general practitioners use computers and hospital doctors do not—Part 1: incentives," *BMJ*, vol. 325, no. 7372, pp. 1086–1089, 2002.

[19] H. Gupta, A. Dastjerdi, S. Ghosh, and A. Buyya, "iFogSim: A toolkit for modeling and simulation of resource management techniques in internet of things, edge and fog computing environments," 2016, http://arxiv.org/abs/1606.02007.

Hindawi
Journal of Sensors
Volume 2018, Article ID 2625195, 15 pages
https://doi.org/10.1155/2018/2625195

Research Article

Beacon-Based Time-Spatial Recognition toward Automatic Daily Care Reporting for Nursing Homes

Tatsuya Morita, Kenta Taki, Manato Fujimoto, Hirohiko Suwa, Yutaka Arakawa, and Keiichi Yasumoto

Graduate School of Information Science, Nara Institute of Science and Technology, Ikoma, Japan

Correspondence should be addressed to Manato Fujimoto; manato@is.naist.jp

Received 22 February 2018; Accepted 4 July 2018; Published 28 August 2018

Academic Editor: Mucheol Kim

As the world's population of senior citizens continues to grow, the burden on the professionals who care for them (carers) is also increasing. In nursing homes, carers often write daily reports to improve the resident's quality of life. However, since each carer needs to simultaneously care for multiple residents, they have difficulty thoroughly recording the activities of residents. In this paper, we address this problem by proposing an automatic daily report generation system that monitors the activities of nursing home residents. The proposed system estimates the multiple locations (areas) at which residents are situated with a BLE beacon, using a variety of methods to analyze the RSSI values of BLE signals, and recognizes the activity of each resident from the estimated area information. The information of the estimated activity of residents is stored in a server with timestamps, and the server automatically generates daily reports based on them. To show the effectiveness of the proposed system, we conducted an experiment for five days with four participants in cooperation with an actual nursing home. We determined the proposed system's effectiveness with the following four evaluations: (1) comparison of performance of different machine-learning algorithms, (2) comparison of smoothing methods, (3) comparison of time windows, and (4) evaluation of generated daily reports. Our evaluations show the most effective combination pattern among 156 patterns to accurately generate daily reports. We conclude that the proposed system has high effectiveness, high usability, and high flexibility.

1. Introduction

In recent years, due to the astonishing progress of medical technology, human life expectancy has shown a consistent tendency to increase [1]. The senior populations continue to increase year by year especially in developed countries. Japan has one of the world's highest life-expectancy rates, and a quarter of its people are over 65. Such aging societies pose many problems for society. One critical problem is the increase of the burden on carers in nursing homes, who are responsible for multiple elderly people, that is, nursing home residents.

According to an interview with the owner of an actual nursing home, Ikoi-no-ie 26 (Ikoi-no-ie 26 (Japanese website): http://www.lifecarejp.com), the burden on the carers at such facilities continues to increase, and many carers leave the industry quickly due to the harshness of their work. Many nursing homes suffer from a chronic shortage of carers. In nursing homes, carers perform a wide variety of tasks, including health checks, rehabilitation support, food preparation, restroom support, and conversation. As illustrated in Figure 1, they also have to write daily reports (that record some primary resident activities) that are used by the government to determine funding for nursing homes. Under aging societies, the situation degrades the service quality, since the carers must prepare daily reports on multiple residents, which disrupt other tasks.

For example, assume a situation where a carer is looking after two residents. One resident is doing her rehabilitation work. The carer has to fill out a daily report on her rehabilitation progress while simultaneously supporting that rehabilitation activity. The other resident is going to the restroom. The carer has to stop filling out the daily reports and support him, which means that the carer cannot accurately report on the rehabilitation. Hence, recording the activities of multiresidents is very hard and writing daily reports places a heavy

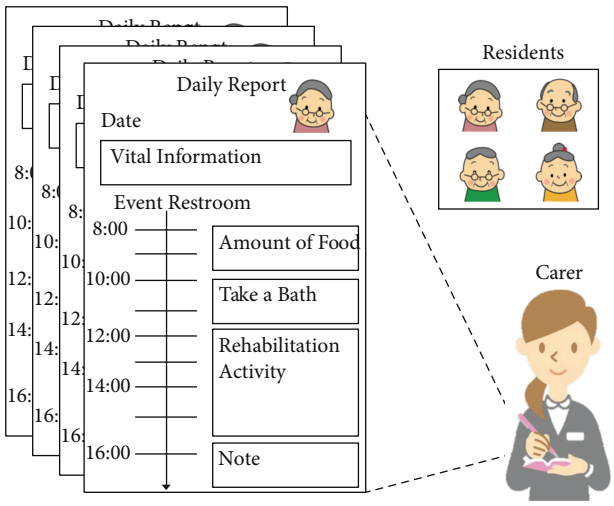


Figure 1: Daily report for each resident.

burden on carers [2]. This burden is a serious problem that impairs the ability of nursing homes to provide safe, high-quality service and must be solved.

To help alleviate this problem, we aimed to determine how ICT/IoT technology can be used to support the recording work of carers in nursing homes. We set the following goals for the desired system: (1) creating a multiresident activity monitoring system and (2) automatically generating daily reports for each resident.

In this paper, in order to achieve these goals, we propose a system that uses BLE beacons, coupled with machine learning using the RSSI values of the BLE signals as the feature quantity, to recognize the location of each resident in the nursing home. The estimated location of each resident is then used to identify the activity they are engaged in. The information of the recognized activity is stored in a server with timestamps, and the server automatically generates daily reports based on them.

Our work makes the following contributions:

(i) First, we provide an architecture called a Movable-Beacon and Fixed-Scanner (MBFS), which is our proposed system's key concept. In this architecture, a BLE beacon is carried by residents and scanners are installed on the target area on the environment side. In existing studies [3–5], many beacons are placed in the environment and users have smartphones for estimating their current area. We call this architecture a Fixed-Beacon and Movable-Scanner (FBMS). However, using FBMS in nursing homes is unrealistic because of low smartphone penetration among residents and the burden caused by forcing them to always carry smartphones. In this paper, we argue that the presented architecture decreases the burden on residents.

(ii) Second, we describe the system based on the MBFS architecture that can monitor the activity of nursing home residents and automatically generate daily reports corresponding to each of them. Specifically, we show that our system is a simple composition which has only four components—a BLE beacon, a scanner, a server, and a web application—and a very useful design that can easily check the information of each resident by carers.

(iii) Third, we develop a BLE beacon with an accelerometer. This BLE beacon consists of the following three components: an accelerometer, a microprocessor with a BLE module, and a battery. A BLE beacon, which is attached to the waists of the residents, has a unique development and resembles a name tag. Therefore, its burden on residents is small. In Section 6, we describe how our developed BLE beacon will be used for rehabilitation support in the future.

(iv) Fourth, we show the effectiveness of our proposed system through a five-day experiment with four participants in cooperation with a nursing home called Ikoi-no-ie 26. We evaluated the effectiveness of our proposed system and expand our previous work [6], by conducting the following four evaluations: (1) comparison of machine-learning algorithms for classification, (2) comparison of smoothing methods, (3) comparison of time windows, and (4) evaluation of generated daily reports. Our evaluations provide the most effective combination pattern in the proposed system to accurately generate daily reports.

(v) Finally, we show that the proposed system clearly has high effectiveness and conclude that the research goals set in this study are sufficiently achieved based on the experimental results.

The rest of this paper is organized as follows. The preliminary work is provided in Section 2, and Section 3 presents our automatic daily report generation system. In Section 4, we describe the experimental environments and evaluation methods. The evaluation results are described in Section 5, and a discussion of the usability and flexibility of our system is provided in Section 6. Section 7 reviews existing work related to this paper. Finally, Section 8 concludes this paper.

2. Preliminary Work

The goals for our study were to (1) create a multiresident activity monitoring system and (2) automatically generate daily reports for each resident. To achieve them, we conducted several preliminary works under the cooperation of Ikoi-no-ie 26. In this section, we explain the basic working environment in nursing homes and their motivation, which is crucial to conduct our study. Then, we describe previous works [2, 7–10] and explain the requirements for the development of our proposed system, as described in Section 3.

2.1. Basic Knowledge and Motivation. In nursing homes, carers perform many kinds of tasks while simultaneously taking care of several residents. For example, they are involved

in health checks, rehabilitation support, food preparation, restroom support, conversations, and so on. Recording daily reports on the residents is one of their most important tasks. The daily report is the nursing report that monitors the activity of residents and records their activity in detail. Such reports are used for the following three main purposes: (1) information sharing among carers, (2) deciding subsequent rehabilitation plans, and (3) communication between residents' families and carers. Hence, daily reports are critical for both residents and carers.

However, from interviewing the nursing home's owner, we found that recording the daily reports of the resident activities is onerous. This problem has also been cited in several papers. Miwa et al. [11] reported that recording work accounts for approximately 25% of the daily work of carers. Inoue et al. [12] described documentation as one of the most time-consuming duties for nursing reports. The recording and documentation responsibilities of daily reports are the biggest barriers to improving the quality of service in nursing homes.

Based on these reasons, we believe that solving the above problems is very important and academically and socially valuable in aging societies. Thus, to lighten the heavy burden of recording the daily reports, we focused on the development of a system that monitors several nursing home residents and automatically generates daily reports corresponding to each of them.

2.2. Previous Work. To realize the above system, it is necessary to monitor a resident in the nursing home. First, as a prototype system, we presented a novel system [7, 8] in which users have BLE beacons and installed scanners on the environment side. In this system, we divided a nursing home into several areas and installed scanners in each one. BLE beacons were attached all day to the clothing of the resident, and scanners received BLE beacon signals. Then by comparing the calculated RSSI values of the BLE signals at each scanner, this system estimated the area with the strongest RSSI value as the resident's current area. As the result of an experiment we conducted at Ikoi-no-ie 26, the existing area of one resident was estimated at a precision of 59.5%. However, the system faced a serious problem. Its area estimation accuracy was too low because the RSSI values dramatically changed based on the installation position of the scanners or the facility's environment. In addition, when installing scanners, we had to consider particular problems of nursing homes, such as residents who are sensitive to changes in their surrounding environment and the positions of the power supply. We needed to camouflage our installed equipment so that residents could not see the scanner positions (Figure 2). Therefore, we needed to place our system at a different height to the installation positions of the scanners or hide them behind training equipment. However, due to the difference in the installation position of scanners, there was the problem where the scanners of neighboring area became nearer than the scanner of the area where the resident was located. Hence, we found that estimating the area where the resident is located is difficult by just simply comparing RSSI values.

(a) Scanner installation work (b) Camouflage

Figure 2: Scanner installation work and camouflage example.

Next, to solve the problem in the above system, we estimated the resident's current area by machine learning [2, 9, 10] with an estimation method that considers the influence of the height difference of each scanner in each area in the nursing home. In our experiment at a nursing home that shows the system's effectiveness, the existing area of one resident is evaluated by the 10-cross-validation method, and the F-measure was 80.6%. In addition, this paper showed that the system can relatively accurately generate daily reports with the estimation results of the existing area. However, we did not conduct any experiments with multiple residents in these papers. Estimating the existing area of multiple residents is important, but actually our experiments were only for one resident. This reason is that it is likely to conflict RSSI signals sent from multiple residents at some BLE scanners and the activity recognition accuracy would be worsened if this system is used as is. Also, these papers pointed out that small erroneous estimations in seconds were generated from the evaluation results. Therefore, we indicated that the accuracy of activity recognition can be further improved by using a smoothing method which is able to eliminate the erroneous estimation.

2.3. System Requirements. Through our previous work and interviews with the nursing home's owner, we identified some problems. Since the nursing home has many residents, our system needs to simultaneously monitor multiple residents. However, our previous work did not consider this. Moreover, to monitor the location of residents, we assumed that they were carrying smartphones that identify their locations. However, the interview with Ikoi-no-ie 26 revealed residents disliked this responsibility since the devices are heavy and bothersome. Therefore, a system using a BLE beacon is more desirable, as in previous work. In addition, a system that does not use cameras or microphones is required in nursing homes because such devices intrude on resident's privacy (We define "privacy intrusion" as "unnecessary and excessive surveillance," e.g., cameras or microphones that monitor residents.)

Thus, we need to develop a system that protects the privacy of residents and develop a low-cost system since most nursing homes are facing budget constraints, that is, insufficient funds to purchase cameras or smartphones for each

resident. In addition, in the nursing home, a system is required that can immediately incorporate new residents because they often change. Based on the above discussion, we define the following system requirements:

Req. 1: it should enable tracking of multiple residents.

Req. 2: it should be a device that places small burden on residents who are carrying it.

Req. 3: it should generally protect residents' privacy.

Req. 4: it should be low-cost.

Req. 5: it should be available for new residents without data collection.

Thus, in this paper, we monitored the activities of multiple residents using a system that satisfies the above five requirements and automatically generated daily reports using the information of each resident.

3. Proposed System

Based on the discussion in Section 2.3, we propose an automatic daily report generation system that monitors the activity of nursing home residents. Our proposed system estimates the present area of multiple residents utilizing machine learning with the RSSI values of BLE signals as the feature quantity and recognizes the activity of each resident from the estimated area information. The information of the recognized activity of each resident is stored in the server with timestamps, and then the server automatically generates daily reports on each resident based on them. In this section, we outline our proposed system, explain the MBFS (Movable-Beacon and Fixed-Scanner) architecture, which is our proposed system's key concept, and describe the application design that is based on the MBFS architecture. Finally, we explain the activity recognition process.

3.1. Outline. In nursing homes, the target activities are area-dependent activities (Figure 3); for example, if a resident is in the restroom area, we can recognize her activity in the restroom. When a resident is in the rehabilitation area, her activity is doing rehabilitation. We installed fixed scanners in the target areas where activities occur that we want to monitor, and residents carry BLE beacons. The system recognizes specific activities by utilizing machine learning with the RSSI values of the BLE signal observed by each scanner as features. In our proposed system, the FBMS (Fixed-Beacon and Movable-Scanner) architecture cannot be adopted in our target environment, since the residents never carry smartphones in the nursing home. Therefore, our proposed system adopts the MBFS architecture through which BLE beacons are carried by residents and scanners are installed in the target area of the environment side. Here, we summarize the following aspects and characteristics of our proposed system:

(i) The proposed system can track the activities of multiple residents because it can distinguish among them by UUID, major value, and minor value included in the BLE signal.

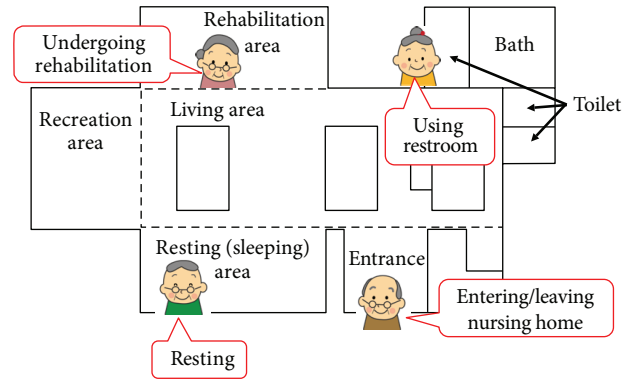


Figure 3: Activities of nursing home.

(ii) The proposed system is based on MBFS architecture. Therefore, its burden on residents is small since they just carry a BLE beacon.

(iii) The proposed system has low privacy concerns because it does not use cameras or microphones. Our system just observes the signals of the BLE beacon.

(iv) The proposed system is relatively inexpensive since its simple composition has only four components: a BLE beacon, a scanner, a server, and a web application.

(v) The proposed system can incorporate new residents without data collection by simply adding a BLE beacon.

Hence, the proposed system satisfies requirements 1 to 5 mentioned in Section 2.3.

3.2. MBFS Architecture. In this paper, we explain an architecture called MBFS to solve the problem of FBMS architecture. Generally, residents never carry smartphones in the nursing home. However, many existing studies estimate their activities using smartphones. In these studies [3–5], many beacons are placed in the environment side and users have smartphones that recognize their activity. We call such architecture FBMS. In fact, when recognizing the activity of nursing home residents, a tag-less system is ideal because little special equipment must be carried. Thus, FBMS architecture, which requires that smartphones be carried, places a heavy burden on residents.

Figure 4 compares the MBFS and FBMS architectures. MBFS architecture is based on an idea that is opposite to FBMS architecture. In recent years, several research studies that resemble MBFS architecture have been conducted [13, 14]. In MBFS architecture, the user side carries a BLE beacon, and scanners are installed on the target areas of the environment side. Therefore, the burden on the residents is very light because they do not need to carry smartphones in the nursing home. BLE beacons are designed to minimize the burden on residents since they are embedded in the name

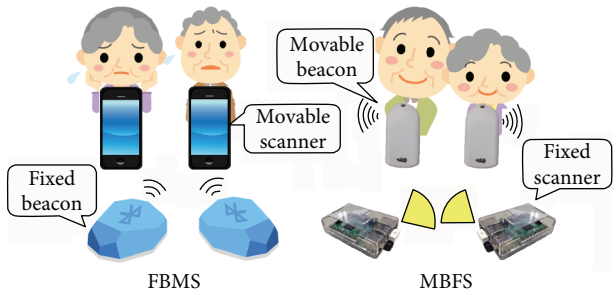


Figure 4: Comparison between FBMS (left) and MBFS (right) architectures.

tags that are constantly worn by the residents or are attached to positions that do not affect daily life, such as the waist. Hence, as a system to be introduced to nursing homes, MBFS architecture is more suitable than FBMS architecture that requires residents to carry smartphones as scanners.

3.3. System Design. In this subsection, we describe the design of the proposed system. Figure 5 shows its composition and data flow. The system is composed of four components: (1) a BLE beacon, (2) a scanner, (3) a server, and (4) a web application. The proposed system's design is based on MBFS architecture. We explain each component below.

(1) *BLE beacon*: the BLE beacon's development is unique, relatively inexpensive, and downsized. Our developed BLE beacon is composed of an accelerometer, a microprocessor with a BLE module, and a battery. It can also send acceleration data in addition to the basic data of such BLE signals as UUID, major value, and minor value. The BLE beacon integrates these data and sends them to each scanner installed on the target areas at intervals of approximately five packets per second. Since our BLE beacon has an accelerometer, it will be used for rehabilitation support in the future. A more detailed explanation of BLE beacons is discussed in Section 6.

(2) *Scanner*: we developed a scanner using a Raspberry Pi, equipped with a BLE dongle (its receiving sensitivity is approximately −100 to −50 dBm) for receiving signals from BLE beacons and a Wi-Fi dongle for sending data to the server. A scanner with Raspberry Pi is less expensive than using a smartphone as the scanner. When each scanner installed in the target areas receives data from the BLE beacons, the signal's RSSI values of each BLE beacon are calculated and the data are sent to the server by Wi-Fi by adding the data of the installation area and a timestamp.

(3) *Server*: the data sent from each scanner are first stored in the server's database (DB). At this time, the server distinguishes each resident by UUID, major value, and minor value. The server estimates the present area of each resident by machine learning, based on the data in the DB. Then the server recognizes the activity of each nursing home resident based on the information of the estimated area of each resident and generates activity reports.

(4) *Web application*: finally, the information on the activity reports of each resident is automatically generated in real time on the web application as daily reports for each resident.

The proposed system is very useful because carers can easily check the information of each resident anytime and anywhere in real time in the nursing home. But of course, this system must accurately recognize the activity of each resident. Thus, in Section 3.4, we describe the activity recognition process in detail.

3.4. Activity Recognition Process. In a small nursing home like Ikoi-no-ie 26, its interior construction is often an open space. Therefore, estimating the current area of each resident is complicated by the influence of the multipath and/or conflict of the BLE signals and the problem of the different height of the installation position of the scanners. In other words, recognizing the activity of each resident is very difficult. Hence, our most important technical challenge is accurately recognizing the activity of each resident. In this subsection, we explain the process of activity recognition in the proposed system. We estimate the present area of residents by utilizing machine learning and recognize the activity of residents based on the information of estimated area. We describe the activity recognition process in the following three phases: (1) data collection, (2) construction of classification model, and (3) smoothing. Each phase is described below.

3.4.1. Data Collection. In our proposed system, we use the RSSI values of BLE signals as features. RSSI values are collected based on the flow described in Section 3.3. Approximately five BLE signals are sent per second. However, BLE signals are not stable because their strength slightly changes depending on the environment. Unstable signals may influence the recognition of the activities of residents. Therefore, before utilizing machine learning, we need to process the data to reduce the influence of unstable signals. We minimize such influence by calculating the simple mean value of the RSSI values that are obtained per second.

Another problem is that packet loss occurs more frequently than expected because BLE signals are comparatively weak. If packet loss occurs, the system might not receive any BLE signals at all. The system cannot accurately recognize the activity of each resident when the RSSI values cannot be obtained. We need to seriously deal with this problem. Thus, we must complement the RSSI values that could not be received in advance by using dummy data, in order to accurately recognize the activity of each resident.

As shown in Figure 6, for the parts (i.e., the parts indicated by "–") where RSSI values could not be obtained due to packet loss, the system complements for them by −100 dBm, which is the lower limit of the receiving sensitivity of a BLE dongle as dummy data. In contrast, for the parts where the RSSI values were obtained, the system calculates the simple mean value of the RSSI values that were obtained per second to reduce the influence of unstable signals. If more

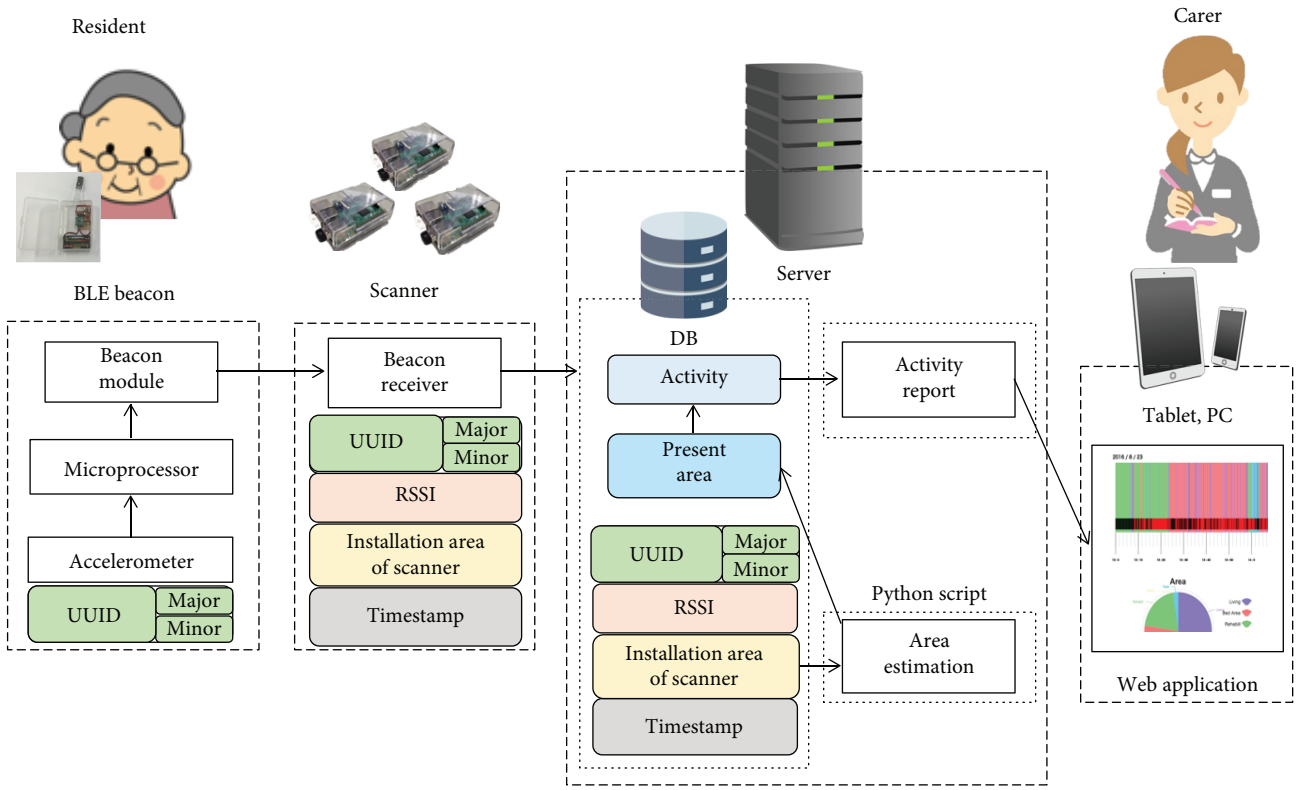


Figure 5: Composition and data flow of the proposed system.

Time	Area 1	Area 2	Area 3
0:00:00	−88 −82 −85 −86	—	−93 −98 −95
0:00:01	−80 −78 −81 −80	−73 −71	—

Time	Area 1	Area 2	Area 3
0:00:00	−85	−100	−95
0:00:01	−80	−72	−100

Figure 6: Data processing in activity recognition process.

than one RSSI value is obtained per second, we do not compensate for the missing values using dummy data even if packet loss occurs. Since the dummy data have very low values compared to the other obtained RSSI values, when simple mean values are calculated using RSSI values including the dummy data, the values become incorrect. In this case, we do not compensate for the missing values using dummy data.

We also installed cameras in Ikoi-no-ie 26 to acquire the groundtruth data after gaining permission from the owner of Ikoi-no-ie 26, its residents, and their families. We manually labeled the collected data while viewing the recorded video. For labeling, we assigned numbers to each area in advance.

3.4.2. Construction of Classification Model. After the data collection, we constructed classification models based on the collected data. We used scikit-learn [15], which is a machine-learning library for constructing the classification models. This library constructs classification models based on a large number of machine-learning algorithms. We constructed classification models by adopting the following four machine-learning algorithms that are used in several kinds of studies [16–19]: (1) logistic regression, (2) support vector machine, (3) random forest, and (4) gradient boosting decision tree.

Logistic regression (LR) [20, 21] is a well-known statistical classification algorithm for predicting dichotomous dependent variables. In addition, we can classify multiple classes by setting multiple applied variables to provide quicker and more robust results than other classification algorithms.

The support vector machine (SVM) [22] is a typical pattern recognition and supervised learning algorithm that analyzes data used for classification and regression analysis. This algorithm's characteristics include a maximum-margin

hyperplane and a kernel trick. It has higher generalization ability among the learning algorithms of data classification.

Random forest (RF) [23] is a machine-learning algorithm based on ensemble learning. It uses multiple decision trees [24] as weak classifiers and obtains classification results by integrating the results from these weak classifiers. It has higher performance with shorter computation time for particular targets than other algorithms.

The gradient boosting decision tree (GBDT) [19, 25] is an ensemble algorithm that uses decision trees. Like other ensemble algorithms, it is built incrementally where each successive estimator reduces the previous model's error. It takes longer to build a model than random forests because each tree has to be built based on the results of a prebuilt tree.

The comparison of the performance of these machine-learning algorithms in our study is described in Section 5.

3.4.3. Smoothing. Figure 7 shows an example of an activity classification result, called a *daily report* in this paper. This figure shows the result of displaying a resident's recognized activities for every second in a time-series order. From this result, we identified many misclassifications, which were caused by the influences of the multipaths and/or conflicts of BLE signals and the problem of the different height of the installation position of the scanners. In this way, even with machine learning, the influence of many unstable signals remains. Therefore, to accurately recognize the activity of residents, we need to reduce the influences of unstable signals even after machine learning. In this paper, we propose the following two smoothing methods: (1) by most frequent value (MFV) and (2) by machine learning.

First, we describe smoothing by the most frequent value (MFV). Figure 8 shows an example of smoothing by the most frequent value (time window is set as odd seconds). We first set a time window and use a range that performs smoothing from the classification result. If the time window is set as odd seconds, we need to give the margin of (time window/2) in the front and rear of the smoothing range. At this time, the decimal point is suppressed. If the time window is set as even seconds, we need to give a margin of {(time window/2)−1} in front of the smoothing range and a margin of (time window/2) in its rear. If the time window is an even number, the way of giving the margin may be reversed. This margin is necessary for counting the most frequent value in the start and end points of the smoothing range, as shown in Figure 8. Even though the data within this margin cannot be displayed on the daily reports, it is not a problem because at most about several tens of seconds of observation data are affected for one day.

Then, we count the most frequent value of the area numbers within the time window. The most frequent value is decided as the classification activity within the time window. Then, we replace the part surrounded by the red squares in Figure 8, which is the middle part of the time window, with the identified classification activity. If there are two or more most frequent values at this time, we determine the classification activity within the time window based on high-priority activity that is set in advance. We defined high-priority activity as activities that we want to reliably recognize in the daily reports, especially the toilet and the bath that contain critical information. To prioritize activities other than toilet and bath, we also decided them comprehensively based on information on the staying time in each area of the residents and advice from the nursing home's owner. We did smoothing while sliding the time window one by one from the start to the end point of the smoothing range.

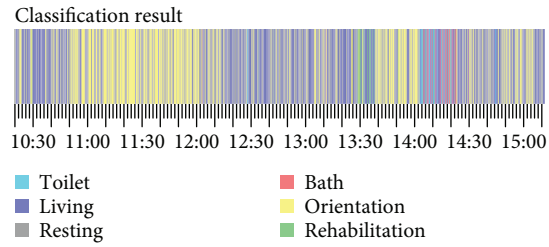


Figure 7: Example of activity classification result (daily report).

Second, we describe smoothing by machine learning. This method smoothens by machine learning using the classification pattern (i.e., the part surrounded by the blue dashed line in Figure 8, e.g., "11514," and "15145") of the activities within the time window as a feature quantity. At this time, the groundtruth data uses the actual activity of a resident. We used three of the four machine-learning algorithms (excluding SVM) described in Section 3.4.2 for creating classification models. For a real system, since tuning based on each nursing home is required, we believe that SVM, which requires learning time, is undesirable in the smoothing process. In this method, in the same smoothing procedure as a smoothing method based on the most frequent value, smoothing is performed using constructed classification models.

A comparison of two kinds of smoothing methods (four patterns) is described in Section 5.

4. Experiment

In this section, we describe our experiment using the proposed system in an actual nursing home. First, we outline our experiment and explain its environment. Finally, we explain the evaluation method.

4.1. Outline of Experiment. To show the effectiveness of our proposed system, we conducted an experiment at Ikoi-no-ie 26 and collected data for five days from four participants (2 males and 2 females). Three of the four participants are residents, and the other is a carer. The data acquired from the carer were used only as training data. As shown in Figure 9, the participants wore our developed BLE beacon on their waist, as this did not obstruct their rehabilitation or recreation activities. The participants then went about their normal day. The data of each participant were collected by the flow described in Section 3.3. Our experiment has the following three purposes: (1) determination of effective machine-learning algorithm, (2) determination of effective smoothing method, and (3) selection of appropriate time window.

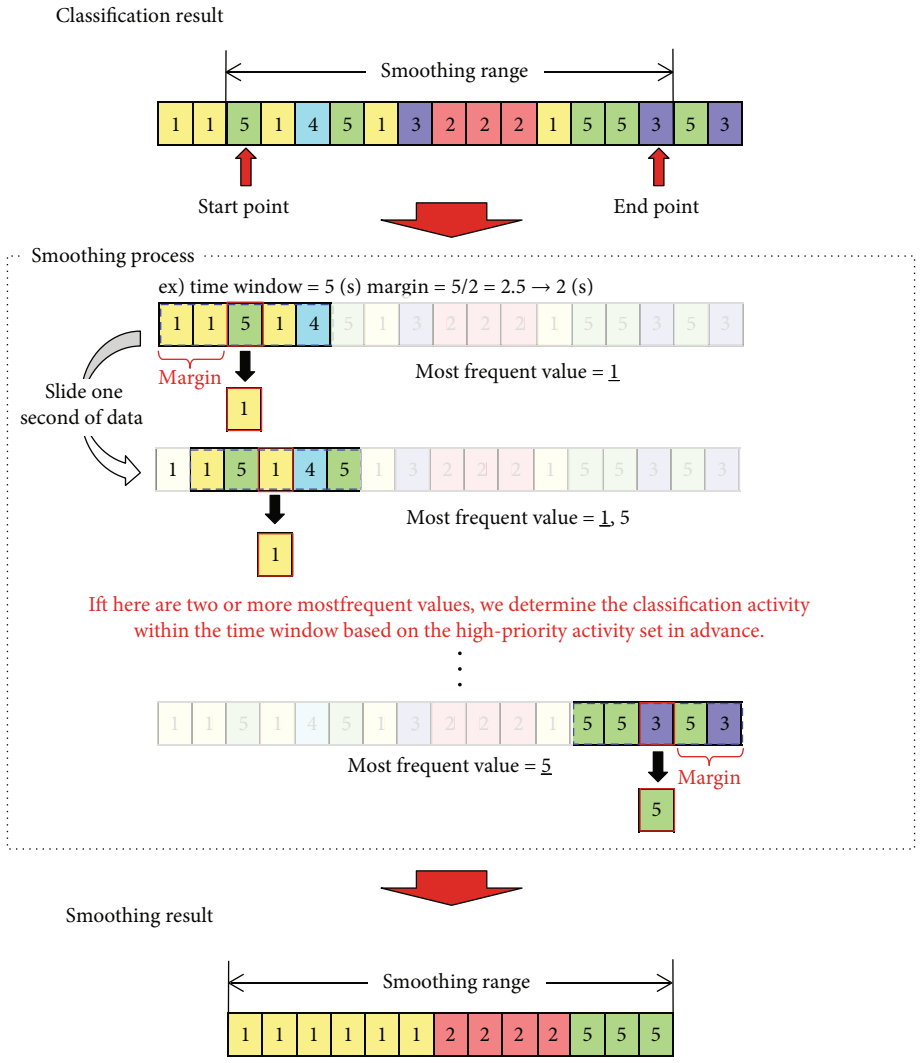


Figure 8: Example of smoothing by most frequent value (time window is set as odd seconds).

Figure 9: BLE beacon attached to waist.

4.2. Environment. Ikoi-no-ie 26 is an actual nursing home that has a 250 m^2 floor. Figure 10 shows its interior view, and Figure 11 shows its floor plan. In the experiment, we divided the floor into these six target areas: toilet, bath, rehabilitation, resting, recreation, and living. We determined them based on discussion with Ikoi-no-ie 26's owner. The toilet and bath areas are closed spaces behind sliding doors, and other areas are open spaces. In the experiment, we tracked the four participants by installing nine scanners at the positions indicated by red circles (Figure 11). A server collected data from scanners and automatically generated activity reports for each participant. Carers can read the daily reports on their tablet or PC. We also installed two cameras to acquire groundtruth data (Figure 11).

4.3. Evaluation Method. We conducted the following four kinds of evaluations and explain their detailed results in Section 5.

(a) *Comparison of Machine-Learning Algorithms for Classification.* We evaluated the classification performance of four kinds of machine-learning algorithms to determine the most effective one. We evaluated the classification performance of the machine-learning algorithms by comparing the recognition accuracy of each activity using leave-one-person-out cross-validation. The important factors in determining an effective machine-learning algorithm for classification are not only high activity recognition

Figure 10: Interior view of Ikoi-no-ie 26.

accuracy but also whether the toilet and bath activities (whose recognition we prioritized in the daily reports) are accurately recognized. Thus, we did not determine an effective machine-learning algorithm for classification by only average recognition accuracy; instead, we comprehensively determined it by investigating the recognition accuracy of each activity.

(b) *Comparison of Smoothing Methods*. We evaluated the performance of two kinds of smoothing methods and determined the most effective one by adapting the smoothing methods (four patterns) to the machine-learning algorithm for the classification determined in evaluation (a). After discussions with Ikoi-no-ie 26's owner, we determined the priority of each activity with machine learning for smoothing as follows: (1) toilet, (2) bath, (3) rehabilitation, (4) resting, (5) recreation, and (6) living.

(c) *Comparison of Time Windows*. We investigated the effect of applying different time windows to the combination pattern determined by evaluations (a) and (b) to select an appropriate time window and determined a final combination pattern from the result.

(d) *Evaluation of Generated Daily Reports*. Finally, we evaluated the effectiveness of the proposed system by comparing the daily reports generated by the groundtruth data and the final combination pattern determined by evaluations (a), (b), and (c). In this evaluation, we showed that the proposed system can incorporate new residents without data collection.

5. Result

5.1. Comparison Result of Machine Learning Algorithms for Classification. First, we evaluated the performance of the machine-learning algorithms for classification by comparing the recognition accuracy of each activity using leave-one-person-out cross-validation. Figure 12 compares the results of the machine-learning algorithms for classification. The horizontal axis shows each activity and the weighted average, and the vertical axis shows the F-measure.

As shown in Figure 12, for the weighted average, all algorithms show a relatively high F-measure over 0.65. SVM, RF, and GBDT keep relatively stable F-measures for all activities. By contrast, LR shows an overall lower F-measure than the above algorithms. In particular, the F-measure for recreation and living activities, which have low priority, is relatively high, but the F-measure for toilet and bath activities, which have high priority, is very low. In addition, the F-measure for rehabilitation and resting activities, which have medium priority, is also very low. From these reasons, we excluded LR with low F-measure in this evaluation. Hence, we determined SVM, RF, and GBDT as effective algorithms for classification in this study.

Figure 13 shows the daily reports (classification results) generated by the determined algorithms (before adapting the smoothing methods). Many misclassifications remain in each daily report that was generated by the selected algorithms. Therefore, we have to reduce them. We believe that daily reports with higher accuracy can be generated by reducing these misclassifications. Thus, we clarified which smoothing methods are effective by applying two kinds of smoothing methods to the determined algorithms in the next subsection.

5.2. Comparison Result of Smoothing Methods. We determined the most effective smoothing method by adapting the smoothing methods (four patterns) to the machine-learning algorithms for the classification determined in Section 5.1. Table 1 shows the comparison result of the smoothing methods (time window = 20 sec). The top row shows the machine-learning algorithms for the classification determined in Section 5.1, and the second row shows the machine-learning algorithms for smoothing. Note that CR shows the classification result before adapting the smoothing methods. Also, the numbers in Table 1 show the F-measure. In this evaluation, we used a value of time window = 20 seconds for all the smoothing methods to confirm the effect of each method.

After adapting the smoothing methods, we found that the F-measure became quite high for all the combination patterns except CR. Therefore, applying smoothing methods is effective. However, when applying machine learning to smoothing, unrecognizable activities occurred. Smoothing by LR cannot recognize (i.e., $F-\text{measure} = 0$) the toilet, rehabilitation, and resting activities, regardless of which machine-learning algorithms we had used for the classification determined. Smoothing by RF or GBDT cannot recognize resting activities. In the daily reports that require detailed information, unrecognized activities obviously cause problems. Thus, machine learning for smoothing is not very effective.

On the other hand, smoothing by the most frequent value (MFV) tends to have a higher F-measure for activities other than resting compared with CR, regardless of the machine-learning algorithms for the classification determined in Section 5.1. The F-measure on resting activities tends to be lower than CR, but this does not mean that this activity

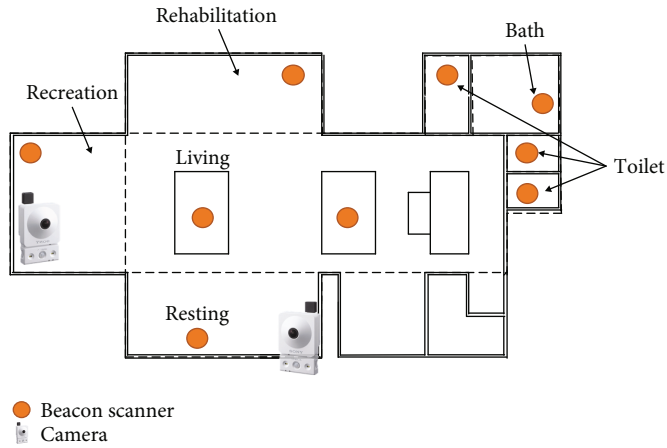


Figure 11: Floor plan of Ikoi-no-ie 26.

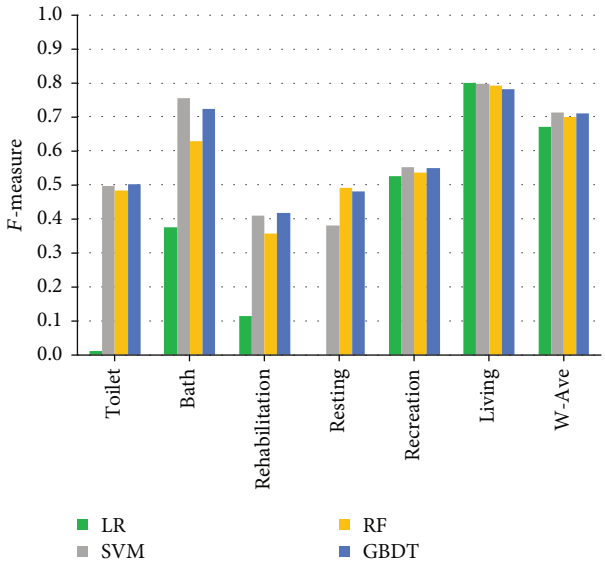


Figure 12: Comparison result of machine learning algorithms for classification.

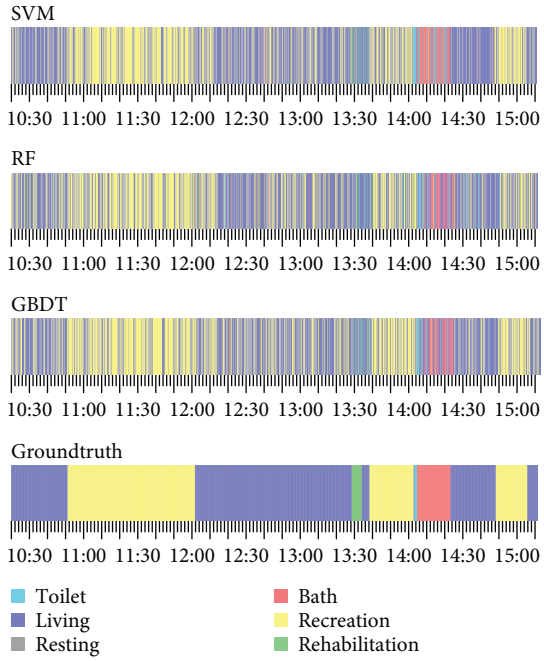


Figure 13: Daily reports generated by determined algorithms (before adapting smoothing methods).

cannot be recognized. Thus, since smoothing by MFV is more effective than smoothing by machine learning, we focus on it.

We compared the combination patterns of RF-MFV, GBDT-MFV, and SVM-MFV. SVM-MFV has a lower F-measure in toilet and bath activities, which have higher priority than other combination patterns. In addition, this combination has a lower F-measure for the weighted average than the other combinations. By contrast, RF-MFV and GBDT-MFV have higher F-measures on toilet and bath activities than SVM-MFV. In daily reports, accurately recognizing information on toilet and bath activities is critical. Hence, in this evaluation, we determined that the combination patterns of RF-MFV and GBDT-MFV have high recognition accuracy of those activities and are the most effective combination patterns among 12 combination patterns.

5.3. Comparison Result of Time Windows. Next, we investigated the effect of applying different time windows to combination patterns determined by Section 5.2 to select an appropriate time window. Figure 14 shows a comparison result of time windows with this combination pattern: RF-MFV, and Figure 15 shows a comparison result of time windows with this combination pattern: GBDT-MFV. The horizontal axis shows the time window, and the vertical axis shows the F-measure. In this evaluation, we investigated the effect of using different time windows of 0 to 120 seconds (13 patterns).

As shown in Figures 14 and 15, we found no significant improvement in the F-measure of each activity even if any time window was used. In such a case, we need to select an appropriate time window based on priority. The most important factor when determining the appropriate time window is

Table 1: Comparison result of smoothing methods (time window = 20 sec).

Activity	RF					GBDT					SVM				
	CR	MFV	RF	GBDT	LR	CR	MFV	RF	GBDT	LR	CR	MFV	RF	GBDT	LR
Toilet	0.483	0.764	0.643	0.725	0	0.501	0.767	0.671	0.743	0	0.497	0.648	0.633	0.688	0
Bath	0.630	0.896	0.872	0.889	0.507	0.723	0.896	0.893	0.903	0.593	0.755	0.885	0.874	0.882	0.748
Rehabilitation	0.355	0.494	0.549	0.616	0	0.416	0.519	0.572	0.638	0	0.408	0.421	0.554	0.604	0
Resting	0.490	0.301	0	0	0	0.480	0.297	0	0	0	0.380	0.105	0	0	0
Recreation	0.535	0.636	0.606	0.628	0.542	0.548	0.647	0.593	0.616	0.527	0.553	0.611	0.557	0.593	0.529
Living	0.793	0.865	0.765	0.764	0.867	0.783	0.862	0.744	0.758	0.863	0.797	0.869	0.721	0.755	0.867
Weighted average	0.700	0.791	0.716	0.725	0.722	0.710	0.793	0.701	0.720	0.720	0.714	0.782	0.675	0.709	0.731

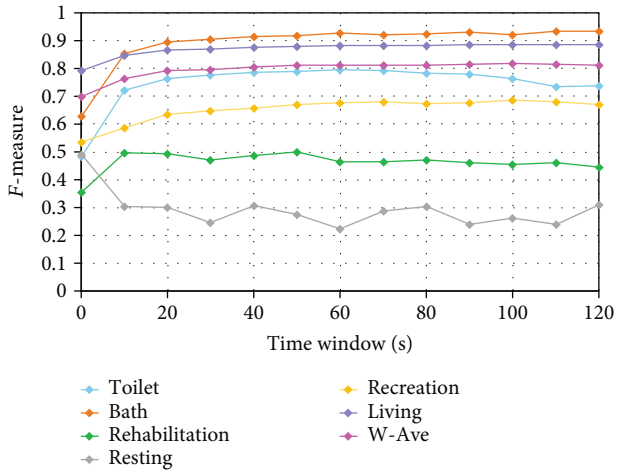


Figure 14: Comparison result of time windows (RF-MFV).

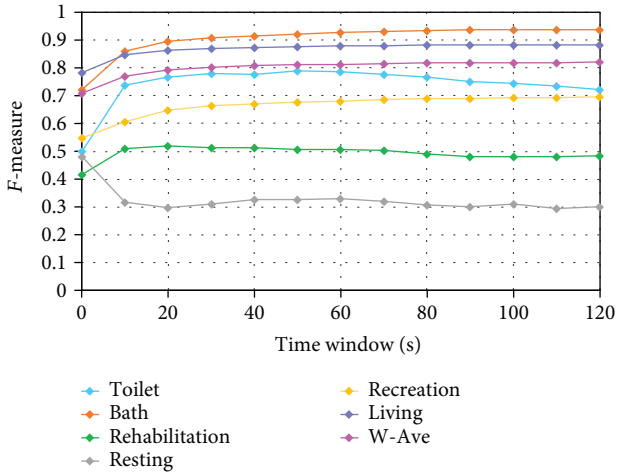


Figure 15: Comparison result of time windows (GBDT-MFV).

that the F-measure is high on the toilet and bath activities that have high priority. Based on such a policy, from Figures 14 and 15, we decided the time window of each combination pattern as follows. The time window is 60 seconds for RF-MFV and 50 seconds for the GBDT-MFV. In this evaluation, we determined these patterns as final combination patterns among 156 combination patterns.

5.4. Evaluation Result of Generated Daily Report. We evaluated the effectiveness of our proposed system by comparing the daily reports generated by the groundtruth data and by the final combination patterns determined by Section 5.3. Figure 16 shows the daily reports generated by the final combination patterns. The upper graph shows a daily report generated by RF-MFV (time window = 60 sec). The middle graph shows a daily report generated by GBDT-MFV (time window = 50 sec). The lower graph shows a daily report generated by the groundtruth data.

From Figure 16, we confirmed that the proposed system can accurately generate daily reports to some extent using RF-MFV (time window = 60 sec) and GBDT-MFV (time w indow = 50 sec). RF-MFV (time window = 60 sec) can basically output toilet, living, and recreation activities. We also found that GBDT-MFV (time window = 50 sec) can exactly output toilet, bath, and recreation activities to some extent. However, we still see far too many misclassifications from the result in Figure 16. Such misclassifications are not serious because we can relatively accurately grasp a participant's activity for a single day. Hence, we conclude that the final combination patterns are sufficiently effective based on these results.

Finally, Table 2 shows the evaluation result for the three participants who are Ikoi-no-ie 26 residents. For the weighted average, the proposed system has high scores that exceed approximately 0.80 in each precision, recall, and F-measure index. Participants A and C have especially high scores for each index. Also, participant B has the high score for each index, although it is somewhat lower than the scores of participants A and C. We achieved these results by evaluating the proposed system with leave-one-person-out cross-validation. In other words, this result means that the proposed system can incorporate new residents without data collection. Hence, it satisfies requirement 5 that we discussed in Section 2.3. This result also shows that the proposed system can accurately track the activity of multiparticipants. It also satisfies requirement 1 in Section 2.3. Hence, our proposed system sufficiently achieved the research goals of this study.

6. Discussion

In the evaluations in Section 5, we determined RF-MFV (time window = 60 sec) and GBDT-MFV (time window =

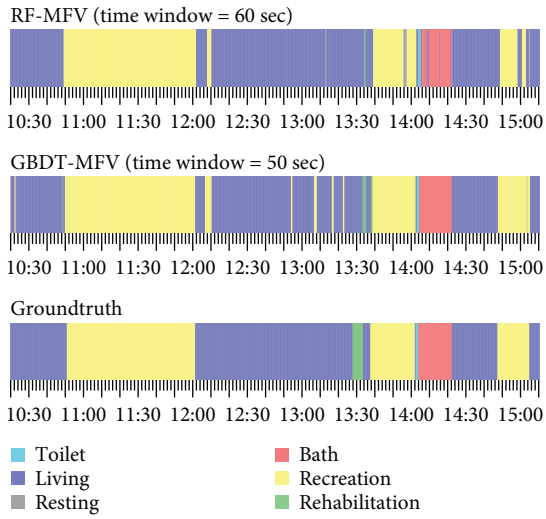


Figure 16: Daily reports generated by final combination patterns.

50 sec) as the most effective among 156 combination patterns. We found that the daily reports generated by these effective combination patterns are accurately output to some extent. The same result was obtained for other residents. These results mean that the proposed system simultaneously monitored the activity of multiresidents and automatically generated daily reports corresponding to each resident. Thus, our proposed system sufficiently achieved our research goals. Also, this system sufficiently satisfied all the requirements described in Section 2.3 from the system design and evaluation results. In this section, we discuss the usability and scalability of our proposed system to derive a clearer conclusion in this paper.

6.1. Usability. First, we discuss the usability of the proposed system. To investigate its usability, we discussed it with the owner of Ikoi-no-ie 26 and interviewed the experiment's participants about the mobility of our BLE beacon. We obtained the following results from the owner and participants:

(i) The proposed system was an adequate prototype for Ikoi-no-ie 26, since the activities of each resident and new resident were visualized without any manual work by carers. It satisfies requirements 1 and 5 in Section 2.3.

(ii) The developed BLE beacon used in the proposed system was light enough for seniors to carry. It satisfies requirement 2 in Section 2.3.

(iii) The proposed system raises no privacy concerns since it does not use cameras or microphones. It satisfies requirement 3 in Section 2.3.

(iv) The proposed system is easy to introduce even in a small nursing home, since it is compact and inexpensive. It satisfies requirements 4 in Section 2.3.

The above results satisfy all requirements in Section 2.3 and demonstrate that our proposed system is sufficiently helpful for the nursing duties/tasks of carers in nursing homes. Furthermore, it places a small burden on users and low privacy concerns for residents. Hence, since it is very easy-to-use for both carers and residents, it has high usability.

6.2. Flexibility. Next, we discuss the flexibility of our proposed system. Especially, we believe the flexibility of our developed BLE beacon is very high. Since our BLE beacon has an accelerometer with a 3-axis accelerometer, a gyroscope, and a geomagnetic sensor, we can obtain various useful information, for example, detailed activities and the number of steps by analyzing the obtained data of each sensor. We believe that utilizing our BLE beacon will be used for rehabilitation support in the future in aging societies. Here, we list three examples that show the flexibility of our BLE beacons:

(i) The developed BLE beacon has a classifier that recognizes the following activities: "sit," "stand," and "walk." It can recognize the kind of activities performed by nursing home residents. More detailed daily reports might be generated by utilizing the results of these activity recognitions [2, 9, 10].

(ii) The developed BLE beacon can measure the number of steps of each resident by analyzing the obtained acceleration data. Perhaps the physical activity of each resident can be grasped by measuring these steps. The physical activity recognition of residents is also important for the daily reports, since the ratio of a day's activity is the key information to obtain the states of their health.

(iii) The developed BLE beacon has a 3-axis accelerometer, a gyroscope, and a geomagnetic sensor. Therefore, perhaps the type of rehabilitation can be recognized by integrating and analyzing these sensor data. If our BLE beacon can recognize the type of rehabilitation, the completeness of the daily reports will be further improved.

Hence, we conclude that our proposed system has high flexibility.

7. Literature Review

There are many studies related to our proposed system. In this section, we discuss them in the following three subsections: (1) location estimation technique, (2) activity recognition technique, and (3) activity report generation in health care.

7.1. Location Estimation Technique. Location estimation techniques have been proposed based on several kinds of techniques, including Wi-Fi signals [26], BLE signals [3, 4, 27], magnetic fields [28], infrared [29], and ultrasounds [30]. For example, camera-based systems are one of the most famous techniques [31, 32]. However, they are not suitable for our target environment because they cannot distinguish people unless their face is captured clearly. If we set several cameras in the corners of the center room, the

Table 2: Evaluation result for participants.

Participant	RF-MFV (TW = 60 sec)			GBDT-MFV (TW = 50 sec)		
	Precision	Recall	F measure	Precision	Recall	F measure
A	0.840	0.816	0.791	0.843	0.821	0.798
B	0.787	0.775	0.751	0.785	0.769	0.742
C	0.890	0.879	0.875	0.887	0.867	0.868
Weighted average	0.842	0.826	0.809	0.842	0.822	0.806

person's face captured by the camera is too small to be recognized. If we set so many cameras, costs are increased, and personal privacy is violated as camera accuracy increases. Thus, they are unsuitable for our target environment.

Location estimation techniques based on passive RFID systems have also been proposed [33–37]. These techniques mainly estimate location using a probabilistic model called a sensor model or the directivity formed by the combination of a reader antenna and an RFID tag. However, since they require at least a reader antenna or a special device to receive radio waves from passive RFID tags, we cannot use them in our target environment because they increase the burden on residents. In addition, these studies did not simultaneously estimate the locations of multiple people or objects.

Location estimation techniques based on Wi-Fi and/or BLE signals have also been proposed [38–42]. These techniques achieved comparatively high localization accuracy by estimating location based on triangulation, fingerprints, signal strength (SS), time difference of arrival (TDOA), time of arrival (TOA), and angle of arrival (AOA). However, they require such devices as smartphones to receive Wi-Fi and/or BLE signals. Therefore, the burden on residents is increased. So, we cannot use them in our target environment.

7.2. Activity Recognition Technique. Some studies focus on activity recognition [17, 43, 44]. Activity recognition techniques that use wearable accelerometers have already achieved accuracies exceeding 90% for such simple activities as walking, sitting, running, and sleeping [45]. Lee and Mase [46] proposed a technique to estimate "walk," "run," and "sit" with an accelerometer and a gyroscope attached to users. The technique of Bao and Intille [47] used five wearable accelerometers and recognized 20 activities, such as watching TV, cleaning, and working. Ueda et al. [48] introduced an activity recognition system in a smart home utilizing an ultrasonic sensor-based indoor positioning system and power meters attached to each home appliance.

To utilize the above techniques, residents have to carry a smartphone or several wearable devices. However, they dislike carrying them, as we described in "location estimation technique." Moreover, most collect acceleration data with smartphones and conduct machine learning using high-performance computers. Thus, these techniques are not suitable for our target environment.

7.3. Activity Report Generation in Health Care. Other studies addressed activity report generation in health care [49–51]. Inoue et al. [12] investigated the activity recognition of nurses in hospitals and estimated their activities with accelerometers by mobile devices carried by nurses and Bayesian estimation. They report that one of the most time-consuming duties is the documentation required for nursing reports. Documentation in both hospitals and nursing homes is a bottleneck that must be removed before service quality can be improved. However, these studies assume that users always carry mobile devices such as smartphones. Furthermore, since these studies cannot respond to new residents, they are unsuitable for our target environment.

8. Conclusion

We proposed an automatic daily report generation system that monitors the activity of residents in nursing homes. The proposed system estimates the current area of multiple residents utilizing machine learning with the RSSI values of BLE signals as a feature quantity and recognizes the activity of each resident from the estimated area information. The information of the recognized activity of each resident is stored in the server with timestamps. The system automatically generates daily reports corresponding to each resident based on them. To show the effectiveness of our proposed system, we conducted an experiment at an actual nursing home. The following are our primary findings:

(i) The RF-MFV (time window = 60 sec) and GBDT-MFV (time window = 50 sec) combination patterns are the most effective among 156 combination patterns.

(ii) The proposed system can accurately generate daily reports that correspond to each resident to some extent using the above combination patterns.

(iii) The proposed system can accurately track the activities of multiple participants.

(iv) The proposed system can incorporate new residents without data collection.

(v) The proposed system has high usability and high flexibility based on the discussion in Section 6.

Since our proposed system satisfactorily achieved our research goals, we conclude that it can sufficiently assist the daily work of carers in nursing homes. On the other hand, the problem of obtaining information (especially food and information about vital signs) that we cannot obtain in our proposed system is critical. Future work will develop a system

that easily obtains such information and encourages input at precise timing to terminals possessed by carers when residents and carers enter an area where the above information must be input.

Data Availability

The data used to support the findings of this study have not been made available because the authors have a contract for the data usage with the nursing home that was used as the experimental facility.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work is partly supported by the Japanese Government Monbukagakusho: Japan Society for the Promotion of Science KAKENHI Grant nos. 16H01721 and 16K00126. Also, the authors would like to thank Mister Yukawa, the owner of Ikoi-no-ie 26, for giving them many attractive advises to promote this work as well as providing Ikoi-no-ie 26 as a testbed for this research.

References

[1] "Life expectancy at birth, total (years)," 2014, http://data.worldbank.org/indicator/SP.DYN.LE00.IN.

[2] M. Fujimoto, "Smart sensing for automatic care report generation in day care center," *2017 International Workshop on Smart Info-Media System in Asia (SISA 2017)*, pp. 337–341, IEICE, Fukuoka, Japan, 2017.

[3] J. Paek, J. G. Ko, and H. Shin, "A measurement study of ble ibeacon and geometric adjustment scheme for indoor location-based mobile applications," *Mobile Information Systems*, vol. 2016, Article ID 8367638, 13 pages, 2016.

[4] D. Chen, K. G. Shin, Y. Jiang, and K.-H. Kim, "Locating and tracking ble beacons with smartphones," in *Proceedings of the 13th International Conference on emerging Networking EXperiments and Technologies - CoNEXT '17*, pp. 263–275, Incheon, Republic of Korea, 2017.

[5] Y. Zhuang, J. Yang, Y. Li, L. Qi, and N. El-Sheimy, "Smartphone-based indoor localization with bluetooth low energy beacons," *Sensors*, vol. 16, no. 5, p. 596, 2016.

[6] T. Morita, K. Taki, M. Fujimoto, H. Suwa, Y. Arakawa, and K. Yasumoto, "Ble beacon-based activity monitoring system toward automatic generation of daily report," in *2018 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, pp. 172–177, Athens, Greece, 2018.

[7] K. Komai, M. Fujimoto, Y. Arakawa, H. Suwa, Y. Kashimoto, and K. Yasumoto, "Beacon-based multi-person activity monitoring system for day care center," in *2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, pp. 1–6, Sydney, NSW, Australia, 2016.

[8] K. Komai, M. Fujimoto, Y. Arakawa, H. Suwa, Y. Kashimoto, and K. Yasumoto, "Elderly person monitoring in day care center using bluetooth low energy," in *2016 10th International Symposium on Medical Information and Communication Technology (ISMICT)*, pp. 1–5, Worcester, MA, USA, 2016.

[9] Y. Kashimoto, T. Morita, M. Fujimoto, Y. Arakawa, H. Suwa, and K. Yasumoto, "Implementation and evaluation of daycare report generation system based on BLE tag," in *Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia - MUM '16*, pp. 327–329, Rovaniemi, Finland, 2016.

[10] Y. Kashimoto, T. Morita, M. Fujimoto, Y. Arakawa, H. Suwa, and K. Yasumoto, "Sensing activities and locations of senior citizens toward automatic daycare report generation," in *2017 IEEE 31st International Conference on Advanced Information Networking and Applications (AINA)*, pp. 174–181, Taipei, Taiwan, 2017.

[11] H. Miwa, T. Fukuhara, and T. Nishimura, "Service process visualization in nursing-care service using state transition model," in *Advances in the Human Side of Service Engineering*, p. 482, CRC Press, 2012.

[12] S. Inoue, N. Ueda, Y. Nohara, and N. Nakashima, "Mobile activity recognition for a whole day: recognizing real nursing activities with big dataset," in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15*, pp. 1269–1280, New York, NY, USA, 2015.

[13] J. G. Lee, J. Kim, S. W. Lee, and Y. W. Ko, "A location tracking system using ble beacon exploiting a double-gaussian filter," *KSII Transactions on Internet and Information Systems*, vol. 11, no. 2, 2017.

[14] K. Urano, K. Hiroi, K. Kaji, and N. Kawaguchi, "A location estimation method using ble tags distributed among participants of a large-scale exhibition," in *Proceedings of the 13th International Conference on Mobile and Ubiquitous Systems: Computing Networking and Services - MOBIQUITOUS 2016*, pp. 124–129, Hiroshima, Japan, 2016.

[15] F. Pedregosa, G. Varoquaux, A. Gramfort et al., "Scikit-learn: machine learning in python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[16] K. Lee and M.-P. Kwan, "Physical activity classification in free-living conditions using smartphone accelerometer data and exploration of predicted results," *Computers, Environment and Urban Systems*, vol. 67, pp. 124–131, 2018.

[17] E. Zdravevski, P. Lameski, V. Trajkovik et al., "Improving activity recognition accuracy in ambient-assisted living systems by automated feature engineering," *IEEE Access*, vol. 5, pp. 5262–5280, 2017.

[18] B. P. Roe, H.-J. Yang, J. Zhu, Y. Liu, I. Stancu, and G. McGregor, "Boosted decision trees as an alternative to artificial neural networks for particle identification," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 543, no. 2-3, pp. 577–584, 2005.

[19] M. Fernández-Delgado, E. Cernadas, S. Barro, and D. Amorim, "Do we need hundreds of classifiers to solve real world classification problems," *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 3133–3181, 2014.

[20] D. W. Hosmer Jr., S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, vol. 398, John Wiley & Sons, 2013.

[21] K. Coussement and D. Van den Poel, "Churn prediction in subscription services: an application of support vector machines while comparing two parameter-selection techniques," *Expert Systems with Applications*, vol. 34, no. 1, pp. 313–327, 2008.

[22] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[23] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[24] J. R. Quinlan, "Induction of decision trees," *Machine Learning*, vol. 1, no. 1, pp. 81–106, 1986.

[25] J. H. Friedman, "Stochastic gradient boosting," *Computational Statistics & Data Analysis*, vol. 38, no. 4, pp. 367–378, 2002.

[26] Q. Li, L. I. Wei, W. Sun, J. Li, and Z. Liu, "Fingerprint and assistant nodes based wi-fi localization in complex indoor environment," *IEEE Access*, vol. 4, pp. 2993–3004, 2016.

[27] R. Faragher and R. Harle, "Location fingerprinting with bluetooth low energy beacons," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 11, pp. 2418–2428, 2015.

[28] K. P. Subbu, B. Gozick, and R. Dantu, "LocateMe: magnetic-fields-based indoor localization using smartphones," *ACM Transactions on Intelligent Systems and Technology*, vol. 4, no. 4, pp. 1–27, 2013.

[29] E. M. Gorostiza, J. L. Lázaro Galilea, F. J. Meca Meca, D. Salido Monzú, F. Espinosa Zapata, and L. Pallarés Puerto, "Infrared sensor system for mobile-robot positioning in intelligent spaces," *Sensors*, vol. 11, no. 5, pp. 5416–5438, 2011.

[30] C. Medina, J. Segura, and Á. de la Torre, "Ultrasound indoor positioning system based on a low-power wireless sensor network providing sub-centimeter accuracy," *Sensors*, vol. 13, no. 3, pp. 3501–3526, 2013.

[31] Y. Kim, H. Shin, Y. Chon, and H. Cha, "Smartphone-based wi-fi tracking system exploiting the rss peak to overcome the rss variance problem," *Pervasive and Mobile Computing*, vol. 9, no. 3, pp. 406–420, 2013.

[32] P. Biber, H. Andreasson, T. Duckett, and A. Schilling, "3D modeling of indoor environments by a mobile robot with a laser scanner and panoramic camera," in *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566)*, pp. 3430–3435, Sendai, Japan, 2004.

[33] K. Bouchard, J. Lapalu, B. Bouchard, and A. Bouzouane, "Precise passive rfid localization for service delivery in smart home," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing - UbiComp '12*, pp. 769–772, Pittsburgh, Pennsylvania, 2012.

[34] T. Wada, T. Hori, M. Fujimoto, K. Mutsuura, and H. Okada, "An adaptive multi-range-sensing method for 3d localization of passive RFID tags," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E95.A, no. 6, pp. 1074–1083, 2012.

[35] M. Fujimoto, T. Wada, A. Inada et al., "Localization of passive RFID tags by using broad-type multi-sensing-range (B-MSR) method," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E95.A, no. 7, pp. 1164–1174, 2012.

[36] M. Fujimoto, T. Wada, A. Inada, K. Mutsuura, and H. Okada, "Swift communication range recognition method for quick and accurate position estimation of passive RFID tags," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E95.A, no. 9, pp. 1596–1605, 2012.

[37] E. Nakamori, D. Tsukuda, M. Fujimoto et al., "A new indoor position estimation method of rfid tags for continuous moving navigation systems," in *2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pp. 1–8, Sydney, NSW, Australia, 2012.

[38] L. Wang, T.-K. Hon, J. D. Reiss, and A. Cavallaro, "Self-localization of ad-hoc arrays using time difference of arrivals," *IEEE Transactions on Signal Processing*, vol. 64, no. 4, pp. 1018–1033, 2016.

[39] S. A. Golden and S. S. Bateman, "Sensor measurements for wi-fi location with emphasis on time-of-arrival ranging," *IEEE Transactions on Mobile Computing*, vol. 6, no. 10, pp. 1185–1198, 2007.

[40] X. Tian, W. Li, Y. Yang, Z. Zhang, and X. Wang, "Optimization of fingerprints reporting strategy for wlan indoor localization," *IEEE Transactions on Mobile Computing*, vol. 17, no. 2, pp. 390–403, 2018.

[41] D. Niculescu and B. Nath, "Ad hoc positioning system (APS) using AOA," in *IEEE INFOCOM 2003. Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE Cat. No.03CH37428)*, pp. 1734–1743, San Francisco, CA, USA, 2003.

[42] C. Xiao, D. Yang, Z. Chen, and G. Tan, "3-D BLE indoor localization based on denoising autoencoder," *IEEE Access*, vol. 5, pp. 12751–12760, 2017.

[43] K. Moriya, E. Nakagawa, M. Fujimoto et al., "Daily living activity recognition with echonet lite appliances and motion sensors," in *2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pp. 437–442, Kona, HI, USA, 2017.

[44] L. Chen, C. D. Nugent, and H. Wang, "A knowledge-driven approach to activity recognition in smart homes," *IEEE Transactions on Knowledge and Data Engineering*, vol. 24, no. 6, pp. 961–974, 2012.

[45] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, pp. 1192–1209, 2013.

[46] S.-W. Lee and K. Mase, "Activity and location recognition using wearable sensors," *IEEE Pervasive Computing*, vol. 1, no. 3, pp. 24–32, 2002.

[47] L. Bao and S. S. Intille, "Pervasive Computing: Second International Conference, PERVASIVE 2004, Linz/Vienna, Austria, April 21–23, 2004, Proceedings," in *Chapter Activity R*, pp. 1–17, Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.

[48] K. Ueda, M. Tamai, and K. Yasumoto, "A method for recognizing living activities in homes using positioning sensor and power meters," in *2015 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, pp. 354–359, St. Louis, MO, USA, 2015.

[49] M. Tentori and J. Favela, "Monitoring behavioral patterns in hospitals through activity-aware computing," in *2008 Second International Conference on Pervasive Computing Technologies for Healthcare*, pp. 173–176, Tampere, Finland, 2008.

[50] V. Osmani, S. Balasubramaniam, and D. Botvich, "Human activity recognition in pervasive health-care: supporting efficient remote collaboration," *Journal of Network and Computer Applications*, vol. 31, no. 4, pp. 628–655, 2008.

[51] S.-C. Kim, Y.-S. Jeong, and S.-O. Park, "RFID-based indoor location tracking to ensure the safety of the elderly in smart home environments," *Personal and Ubiquitous Computing*, vol. 17, no. 8, pp. 1699–1707, 2013.

Hindawi
Journal of Sensors
Volume 2018, Article ID 6272793, 12 pages
https://doi.org/10.1155/2018/6272793

Research Article

OperaBLE: An IoT-Based Wearable to Improve Efficiency and Smart Worker Care Services in Industry 4.0

Luis Roda-Sanchez,[1] **Celia Garrido-Hidalgo**,[1] **Diego Hortelano**,[1] **Teresa Olivares**,[2] **and M. Carmen Ruiz**[2]

[1]*Albacete Research Institute of Informatics (I3A), Albacete 02071, Spain*
[2]*University of Castilla-La Mancha (UCLM), Albacete 02071, Spain*

Correspondence should be addressed to Diego Hortelano; diego.hortelano@uclm.es

Received 23 February 2018; Revised 17 June 2018; Accepted 9 July 2018; Published 26 August 2018

Academic Editor: Mucheol Kim

Industry 4.0 is leading the Fourth Industrial Revolution transforming traditional factories into smart factories governed by the Internet of Things (IoT). In order to assist smart factory employees, this paper introduces OperaBLE, a Bluetooth Low Energy (BLE) wearable proposal which is aimed at enhancing working conditions and efficiency in Industry 4.0 scenarios. We have developed two innovative algorithms for OperaBLE focused on power awareness as the key-enabling attribute towards success: Low-Frequency Movement Characterisation Algorithm (LoMoCA) and Adaptive Heart Rate Algorithm (AHRA). Novel experiments have been carried out using OperaBLE to determine its operability, reliability, and lifespan. Results obtained during experimentation demonstrate how OperaBLE empowers human-machine collaboration embedding workers in closed-loop performance and ensuring nonharmful working conditions by means of power-aware algorithms. OperaBLE is due to bring digitalisation into smart factories, playing an essential role in the emerging wearable revolution to arise in the following years towards smart production systems.

1. Introduction

The Fourth Industrial Revolution has arisen to transform the current industry model and to introduce digitalisation into traditional factories improving production rates and promoting collaboration. This emerging revolution is the widely known Industry 4.0 [1]. The 4.0 attribute focuses on the Internet of Things (IoT) [2] applied to industrial systems in order to interconnect objects, machines, and humans in smart factories. Collaborative tasks are encouraged by means of increasing production rates and minimising costs. This is achieved by introducing cyberphysical systems (CPS) [3], perceived as the embeddedness of sensors to collect data and perform cognitive algorithms through the Internet.

Sustainability is one of the greatest concerns for this emergent paradigm. In a few years, factories will be immersed into IoT ecosystems composed of heterogeneous networks, which is the reason why power-aware alternatives are mandatory to avoid an additional energy waste. Not only environmental but also social sustainability is encouraged to face the greatest challenges of Industry 4.0. New CPS are due to be transformed into human-centric architectures focused on employees which ensure security, wellness, and comfort towards optimal working conditions and production rates.

The recent proliferation of the wearable market is a key indicator of global success based on the embeddedness of sensors and actuators to enable the development of smarter devices. On one hand, technologies such as Bluetooth Low Energy (BLE) have taken the lead in terms of IoT standardization due to the necessity of greener and more flexible alternatives. Particularly, BLE is leading an emergent industrial transformation due to the recent adoption of BLE-mesh topology, where broadcast capability will play an important role. On the other hand, accelerometers are setting the beginning of a new IoT revolution enabling gesture-based systems to enable a better interaction between humans and machines.

These devices generate huge amounts of data which are due to be processed and filtered to extract specific information, which is a global challenge widely known as big data.

OperaBLE is introduced in this paper in order to face the major challenges of Industry 4.0. This BLE-based prototype of work band is intended to bring digitalisation into smart factories and to enhance working conditions by immersing employees into CPS, taking into account human care services for workers.

It consists of a modular wearable that improves security in industrial environments, being able to predict harmful situations or even working accidents. This low-power wearable will impact positively on the industrial paradigm due to its wide range of novel applications for the Industrial Internet of Things (IIoT). Being focused on movement characterisation techniques and heart rate measurements, OperaBLE uses learning skills to track operators in order to improve the industry value chain from a power-aware perspective.

The novelty of this paper lies on the introduction of adaptive algorithms specifically developed for OperaBLE, which play an essential role in its applicability to the incoming industrial paradigm. In addition, the experiments carried out during this investigation permitted a technical evaluation of OperaBLE according to user performance, success rates, and energy consumption with promising results. The rest of this paper is organized as follows: Section 2 highlights related works. Section 3 describes OperaBLE, focusing on hardware in Section 3.1 and detailing the algorithms developed in Sections 3.2 (LoMoCA) and 3.3 (AHRA). Section 4 introduces the experiments carried out and discusses the results. Eventually, Section 5 offers the remarkable conclusions of this research and proposes future works.

2. Related Works

Accelerometer-based wearables are specially used to keep a track of information about personal activities and their energy consumption. The number of commercial wearables (usually wristbands identified as *smart bands*) has increased in last years. Regarding academic and research works, they can be classified in different topics: medical and rehabilitation, movement recognition, older adults, sports, positioning, and work.

Concerning medical aspects, a complete system to avoid the hospital acquired infection is proposed in [4], using sensors and wristbands equipped with accelerometers. In a similar way, other smart bands are proposed for muscular movement recognition and physiological signals [5], for Parkinson's disease detection [6] and for helping patients with obstructive sleep apnea [7]. Additionally, accelerometer-based devices are widely used for rehabilitation applications [8].

Gesture and movement recognition is a wide research field concerning devices based on accelerometers. This recognition can be performed by wearables [9] or even by smartphones [10]. It uses accelerometers to assist older adults; in [11], authors present an accelerometer for movement monitoring; in [12], a method based on accelerometers for identification of declining balance is proposed; finally, accelerometers are used to detect some predefined movements (stand-to-sit and sit-to-stand) in young and old adults [13]. Furthermore, in [14], the authors design and implement an embedded system to estimate the pedestrian walking position.

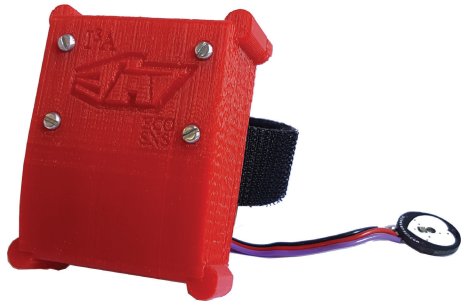
Finally, regarding the use of accelerometers at workplaces, most of articles are focused on the construction sector to enhance security at work. In [15], commercial accelerometers are used to measure the trunk flexion and lateral bending angle in laboratory conditions; this experiment is repeated in a real construction environment in [16]. Both studies are focused on the analysis of the data collected using commercial accelerometers. In [17], an approach based on supervised motion tensor decomposing for construction workers is presented. In [18], authors use the integrated sensor of mobile phones for activity recognition of construction workers. However, other studies are focused on different workers; for example, in [19], authors explore the utility of accelerometers for determining sitting time in office workers. Despite the fact that the use of accelerometers has been evaluated for different professionals, it has not been taken into account for industrial environments to the date, where it could highly contribute to the new Industry 4.0.

Table 1: Principal attributes of OperaBLE.

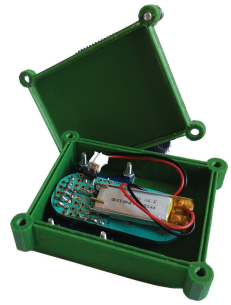
Attributes	Description
Zero fails	Highly reliable algorithms
Sustainability	Power-aware adaptive algorithms
Cooperation	Tap-based human-machine interaction
Standardization	BLE connectivity
Human in the loop	Feedback on worker performance
Health	Human well-being measurements
Intelligence	Movement characterisation
Nearable	Indoor proximity detection
Collaboration	Plug-and-play solution for BLE mesh
Productivity	Supervision for enhancing productivity

3. OperaBLE: The Proposed Wearable

This investigation introduces a wearable intended to satisfy IoT requirements in smart factories. The main attributes of the wearable prototype, OperaBLE, are listed in Table 1. It is an autonomous device able to learn behavioural patterns from individuals and, thus, enhance their working conditions. Although a specific hardware is proposed in this work to implement an evaluable wearable, the major



(a) Exterior design: OperaBLE with pulse sensor



(b) Interior design: OperaBLE with accelerometer

Figure 1: OperaBLE, prototype used for experimentation.

contributions of OperaBLE lie in its functionality and, thus, in the proposed algorithms.

For the development of OperaBLE, different hierarchies have been established, being the first one its reliability for Industry 4.0 applications and the second its sustainable operation. Thus, a right functionality is prioritized over energy efficiency in case of conflict. The key challenge behind OperaBLE is to enable the implementation of its functionality with a high degree of reliability in any lightweight IoT device. This promises to save as much energy as possible due to the reduced microcontroller operating frequency that characterises most tiny IoT devices, which is complimented with the appropriate management of sensors and communication modules.

A technical description of the hardware used to implement the functionality of OperaBLE is provided in Section 3.1. The development of a low-frequency movement characterisation algorithm called LoMoCA, a remarkable contribution of this paper, is addressed in Section 3.2. Moreover, a sustainable algorithm able to adapt the heart rate sampling frequency to each worker, AHRA, is introduced in Section 3.3. Again, the major achievements of this paper are based on the power-aware approach used to develop the mentioned algorithms, which enable the functionality proposed to be run by in lightweight IoT devices typically operating with reduced microcontroller frequencies.

3.1. Hardware Components. This section describes the hardware used to develop the OperaBLE prototype. The core system is the controller board which has an integrated BLE radio module used to transmit data. The device used is LightBlue Bean [20], a low-frequency board that includes I2C and SPI interfaces combined with multiple I/O pins. Regarding integrated sensors, it integrates a temperature sensor and a three-axis accelerometer. Finally, a low-frequency microcontroller, working at 8 MHz, is responsible for managing all parts.

Despite having a three-axis accelerometer integrated in the microcontroller platform, we have assembled an inertial measurement unit (IMU). We included this new board because it has ten degrees of freedom: three-axis accelerometer, three-axis gyroscopes, three-axis magnetometer, and a pressure sensor. Although in this paper we use only the accelerometer and the gyroscope, it enables the use of extra sensors that can be used for further investigation. In addition, a pulse sensor has been included in our prototype to measure the heart rate of operators and to identify harmful situations for them. Eventually, as envelope for the prototype, we have designed a housing using 3D-printing techniques, modelled with the shape shown in Figures 1(a) and 1(b).

One of the main attributes of OperaBLE prototype is its modularity, which is the reason why functionalities to be developed should be tested under different conditions. Our case studio was carried out using two different OperaBLE designs (see Figure 1) for testing individually the two algorithms developed. Considering these characteristics, OperaBLE is a combination of low-frequency and low-cost prototyping board for developing algorithms based on adaptability, low power operation, and sampling frequency.

3.2. Movement Characterisation. This section describes in detail the algorithmic development that permits OperaBLE-characterised movements. In the first place, it is necessary to underline that movement recognition based on accelerometers is not a trivial task, since they provide raw acceleration data that has to be properly filtered and interpreted. For instance, a high difficulty is noticed while separating accurately the body movements and gestures from gravity, for which data sensed by accelerometers is expected to be processed by demanding algorithms that require exhaustive processing and analytics.

As mentioned previously, one of the key achievements of this investigation is a system development for movement recognition using a sampling frequency around 10 Hz, which is up to ten times lower than the one used in novel studies. This fact provides great advantages towards a more efficient battery usage, being highlighted in the following lines the most remarkable contributions:

(1) The movement characterisation algorithm can be implemented in a wide range of devices, including those governed by low-frequency microcontrollers which are less energy-demanding than high-frequency computational platforms [21]. Thus, the minimisation of resources required by this algorithm to process raw acceleration data and identify specific body movements enables its implementation in lightweight and cost-effective devices, contributing towards a more sustainable use of resources.

(2) The algorithm itself, apart from being able to achieve a successful characterisation with a low number of samples per movement, requires as a result a low measurement duty cycle. Downsampling is not an essential energy-saving source of OperaBLE but contributes to the minimisation of energy consumed by accelerometers through the reduction of data acquired.

(3) The use of BLE technology to perform communication enables the optimisation of energy consumed by the communication module as well. A fundamental idea behind the movement characterisation algorithm is that, while data being acquired is not of interest for OperaBLE (which preprocesses raw acceleration data), no information is transmitted to the processing stage. What BLE enables is the possibility to perform sleeping periods between data transmissions with negligible consumption and, as a result of this preprocessing proposed, the amount of data being encapsulated and sent via BLE is greatly reduced.

Although the reduction of either sensor duty cycle or BLE transmissions enhances the energy efficiency of the device, again, the most remarkable contribution to battery saving is provided by the reduced microcontroller frequency demanded by the algorithm. To satisfy the reliability challenge with this low-frequency restriction, we propose to increase the number of different data sources to be analysed. Considering that accelerometers provide data in each space direction, the addition of gyroscopes provides a total of 6 degrees of freedom (DOF). Therefore, we obtain more data to enhance the reliability of the system while low power consumption can be noticed thanks to a low-frequency microcontroller operation required by devices where this algorithm is expected to be implemented.

3.2.1. LoMoCA: Low-Frequency Movement Characterisation Algorithm. This section expands the procedures performed by our algorithm, LoMoCA (see the flow diagram shown in Figure 2 including representative stages). The algorithm of OperaBLE is divided into two parts, differentiated as implementation in a wearable device (preprocessing) and in a nonwearable device (processing).

Regarding the wearable device, we have developed an optimised code to save battery that maintains the device in sleep mode while no significant movement is noticed as can be seen in Figure 2(1). By the time OperaBLE identifies the beginning of a new movement, the subroutine acquires data from the accelerometer and gyroscope integrated in the IMU and stores them in the device until the end of the movement. This movement start event is triggered by a threshold, comparing constantly the current and the previous sample to verify if a difference greater than $1\,m/s^2$ (acceleration) or 1 rad/s (angular velocity) takes place. To determine the end of each movement, the device enters on standby mode while waiting for a significant acceleration or angular velocity during a second. If no movement is perceived during this second, the data collection procedure is considered to be finished.

As stated previously, the acquisition of movement data continues until the previous condition stops being satisfied. Then, the device has two data sources which are checked: accelerometer data (3 DOF) or accelerometer and gyroscope data (6 DOF). The use of 6 DOF considers both angular velocity and acceleration data, which are encoded and sent to the nonwearable device when relevant, see Figure 2(2).

The processing in nonwearable device begins with the data reception event. By the time the analysis stage in Figure 2(3) starts, OperaBLE checks the mode of operation selected in advance. After that, data are decoded properly to proceed to the movement recognition stage. LoMoCA uses a method to recognise movements based on patterns. Thanks to its flexibility, it allows the introduction of new patterns depending on the kind of industry or task that the operator performs. In the first stage, the program checks all patterns stored in a nonwearable device to take into account all of them and compare them with the movement. If the user needed to store a new movement, OperaBLE would be prepared to store it without requiring a reset since the movement recognition procedure also considers patterns recently stored.

At this point, the next analysis procedure is conducted once for each axis and, thus, for each pattern. The most crucial stage is the so-called normalisation of movements (identified in Figure 2(4)). For this, it is necessary to refer the dynamic time warping (DTW) algorithm, since LoMoCA is based on it to evaluate similarity between each pattern and the movement executed. This recursive algorithm was introduced in [22] with the aim of improving existing speech recognition procedures. It is widely used nowadays and is based on the principle of comparing two time series instead of comparing two data sets according to their Euclidean distance (sample by sample). DTW calculates the so-called distance matrix [23], which considers either the comparison of a pair of samples belonging to a time series as well as the adjacent values in order to reach the smallest path and, consequently, the most effective alignment of two time series. The library used during the experiments for applying the DTW algorithm can be accessed in [24]. However, since we noticed that the accuracy of DTW was not enough to achieve a proper movement characterisation with low-frequency sampling, we developed a proprietary algorithm based on DTW, where several issues noticed during the investigation were solved.

In order to carry out the normalisation procedure, axes of patterns are classified in advance. This classification determines which axes are relevant or not in the movement, and then, normalisation plays an important role since DTW is not useful to compare sequences with different amplitudes. Regarding normalisation, for each axis, if the amplitude of acceleration or angular velocity is higher than $4\,m/s^2$ or 4 rad/s, respectively, this axis is normalised using the maximum absolute value in the entire sample. Once this procedure is performed for every axis received, the algorithm compares the concordance between pattern and the normalised movement. Firstly, if the normalisation of the pattern and the movement does not match, the algorithm determines that the incoming movement does not coincide with that pattern. In

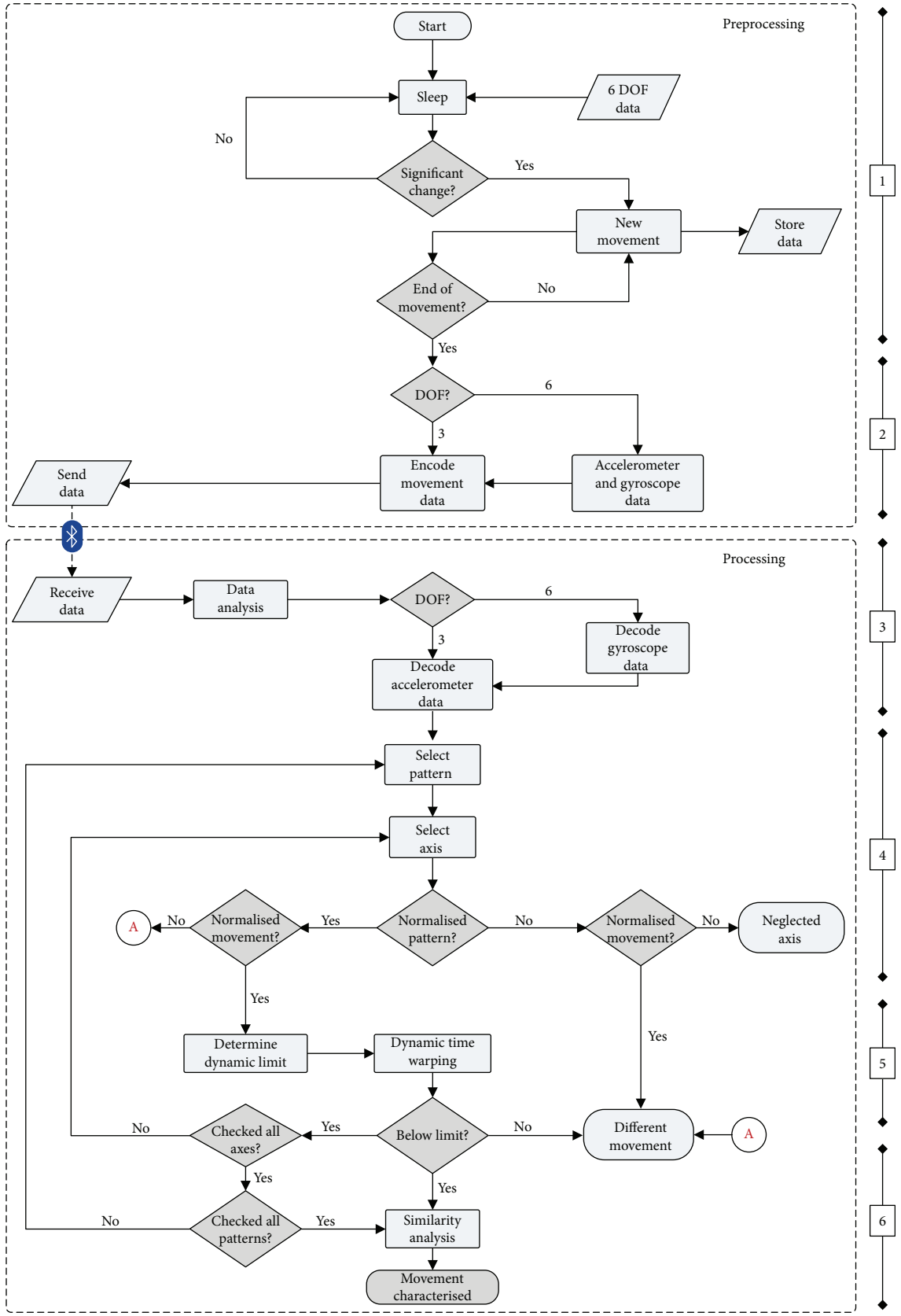


Figure 2: Processing stages of LoMoCA.

the case where pattern and movement are not normalised, the algorithm considers that this axis is not relevant in the movement and discards it for the final decision. Conversely, when the same axes in the pattern and in the sample result to be normalised, the sample is assumed to be a clear candidate to match with the pattern stored. When all axes are analysed in terms of normalisation, the characterisation can continue to the next stage, the dynamic limit determination.

In the next stage (in Figure 2(5)), the dynamic limit is determined according to different parameters such as the length of movement and pattern. For this, an experimental study was conducted to select a dynamic limit which represented similarity between patterns and samples. Equation (1) shows the experimental setting of the parameters used to check similarity, where the threshold is adjusted in each situation depending on the number of samples per movement and pattern considered (N). As can be observed in (1), the limit reaches a constant value for data sets composed of more than 15 samples. Then, DTW is applied to normalised movements and obtain a value of similarity. If all relevant axes in each movement-pattern pair are below the dynamic limit established, a potential compatibility has been found.

$$\text{Limit}_{\text{dynamic}} = 38 \cdot N + 30. \quad (1)$$

Eventually, when all patterns have been checked (in Figure 2(6)), a global similarity analysis is performed to select (from all movement patterns) the closest match and the movement is characterised. For this, the sum of all differential factors obtained (one per each axis analysed in an entire pattern) is taken into account and it is compared with the resulting sums of the other patterns. To finish characterisation, LoMoCA selects the smallest sum as the better pattern-movement match, since the differential factor decreases when similarity increases.

This work provides experiments to discuss the advantages and disadvantages of using accelerometers combined with gyroscopes over using only accelerometers, since a higher energy is typically demanded by gyroscopes. However, the major focus of this section lies in enabling the implementation of the movement characterisation algorithm in lightweight devices with low-frequency microcontrollers, contributing to power-aware approaches.

LoMoCA is a complex algorithm because of the large number of variables to be managed. However, since it is implemented using several filters accurately adjusted, it does not require high computational power. It is important to highlight that LoMoCA is designed for characterising long and short movements, being the last ones is a challenge for most existing movement characterisation algorithms due to the reduced number of samples. OperaBLE is transparent to end users and, hence, ideal to overcome the existing technological gap between digitalised and nondigitalised users. Following this idea, the system implemented is able to learn new movements with no need to modify the software or make changes in the characterisation algorithm, which is an important attribute of OperaBLE.

3.3. Well-Being and Security of Workers. Human-centred architectures are being promoted by means of focusing on the social dimension of sustainability. Apart from the ability for movement characterisation using our low-frequency algorithm, OperaBLE is responsible for assuring nonharmful working environments for factory workers, guaranteeing the prevention of human damages.

OperaBLE has been developed to be embedded in body area networks (BAN) composed of cognitive clothes. Since the focus of this paper is the OperaBLE wearable, the complete BAN is framed by previous studies [25].

In terms of human well-being, OperaBLE measures heart rate in order to monitor unexpected changes with respect to a reference value established during calibration for each worker. We introduce in this work a new algorithm which optimizes the measurement procedure and, thus, saves energy towards greener wearables. The main attribute of AHRA is its capability to adapt the sampling frequency to the state of each worker to detect unhealthy conditions which could lead to accidents.

3.3.1. AHRA: Adaptive Heart Rate Algorithm. This section introduces our adaptive heart rate measurement algorithm, AHRA. Since several existing algorithms achieve high-precision measurements at the expense of low energy efficiency, we developed a proprietary algorithm which is able to adapt sampling frequency by performing flexible sleeping periods which does not affect the quality of measurements.

Initially, the algorithm has a one-minute calibration stage for each subject that requires him or her to be quiet and seated to establish a baseline beat-per-minute (bpm) data. The calibration is only performed once per each worker storing the reference value in EEPROM for daily usage. The algorithm identifies voltage peaks and reaches an average peak-to-peak value of time during a 10-second time measurement. Figure 3 shows a flow diagram of AHRA where low-level routines such as calibration or peak determination are omitted.

According to Figure 3, the reference heart rate value obtained is subsequently used for several analyses and, unless the device is hard reset, additional calibrations will not be necessary. The default interval between measurements is initially set to 30 seconds. If the newest value differs from the reference one, AHRA will question whether it has increased or decreased. In case of decrease, the previous interval is broadened by means of a linear equation (factor B in (3)). Since increases in heart rate must be carefully analysed, in those cases a proportional reduction of time between measurements is established according to a polynomial equation reached by experimentation (factor A in (2)). Factor A was determined considering fixed sampling intervals for specific bpm increases; for instance, a 49-second interval was defined for bpm values equal to the base value calibrated while a 2-second interval was defined for a 10 bpm difference. Considering a set of fixed intervals, the experimental equations (2) and (3) were reached. Furthermore, warning and alert states are defined in AHRA (corresponding to 7% and 14% increases in the reference bpm).

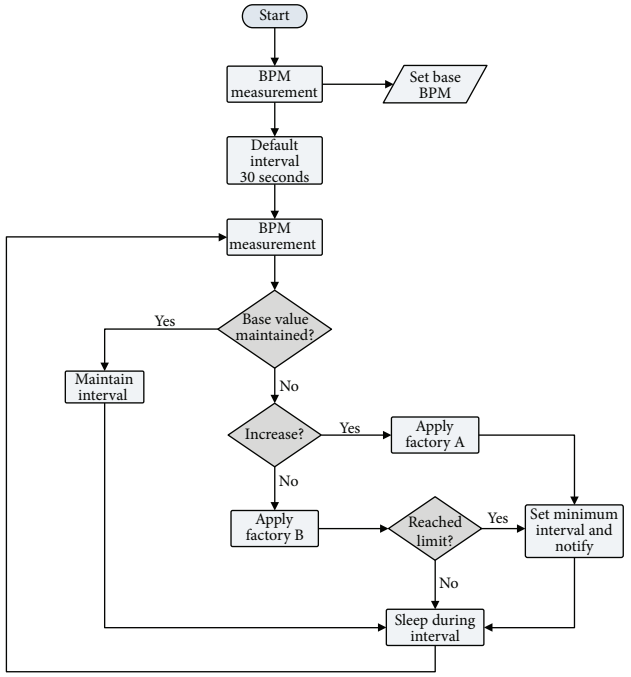


Figure 3: Adaptive heart rate algorithm (AHRA).

$$F_A = -0.140 \cdot (\text{bpm} - \text{bpm}_{\text{base}})^2 - 3.315 \cdot (\text{bpm} - \text{bpm}_{\text{base}}) + 49, \quad (2)$$

$$F_B = (\text{bpm}_{\text{base}} - \text{bpm}) + 50. \quad (3)$$

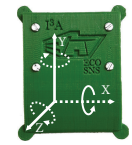
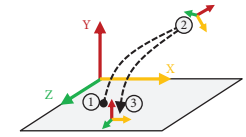
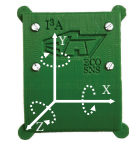
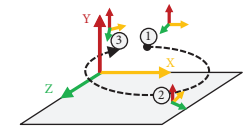
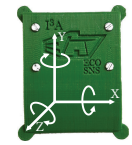
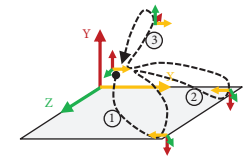
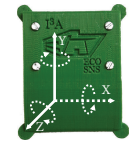
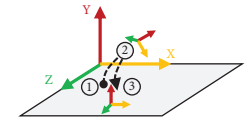
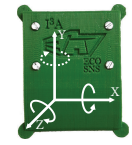
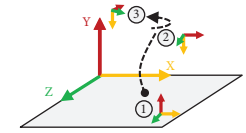
4. Experimental Evaluation and Discussion

This section addresses a set of experiments carried out using OperaBLE and evaluates the energetic contribution of each algorithm. For carrying out the experiments in OperaBLE, a computing board was used as an external processing device to perform the nonwearable role in a LoMoCA algorithm. The load current was measured to evaluate consumption in different conditions using a shunt resistor and an operational amplifier with the aim of reaching a device lifespan comparison.

In this section, several experiments have been defined to evaluate each algorithm individually and reach their major weaknesses. Section 4.1 is focused on LoMoCA, where five different movements were defined to find singularities depending on the number of planes and rotational axes used during characterisation. Section 4.2 explores the performance of AHRA, and for both algorithms, a brief experiment based on power consumption evaluates the suitability of our proposal.

4.1. LoMoCA Experiments. This section defines five baseline movements related to industry-based tasks, to check as many variants as possible and evaluate the system. Table 2 shows a graph including relevant axes in each case and a simplified trajectory representation. These movements have been classified into two groups: the first one includes the so-called *hammer*, *valve opening* and *assembly* (movements thought to represent common industrial tasks for testing

Table 2: Movement patterns for experimentation.

Movement	Relevant axes	Trajectory
Hammer		
Valve opening		
Assembly		
Data request		
Notification		

the algorithm with a progressive difficulty increase) and the second includes *data request* and *notification* (which provide support to help operators in their daily work). All movements were recorded as short movements to have a low number of samples and, thus, hinder the characterisation process. A detailed explanation is provided below:

(i) *Hammer*: uses mainly two axes and only one rotation axis, the 1-2-3 sequence is repeated three times.

(ii) *Valve opening*: this movement was performed using the OperaBLE XY plane. It deals with an additional difficulty because of the centrifugal acceleration that produces a variation of Z axis hindering the recognition procedure. The sequence 1-2-3 is executed four times.

(iii) *Assembly*: consists of a connection task where 6 axes are relevant (most complex movement), which is useful to check LoMoCA accuracy to distinguish among similar movements. The correct sequence is 1-2-3.

(iv) *Data request*: is a special movement recorded to allow operators to request data from the machines using a straightforward method, three taps on the machine surface. It avoids the use of screens, buttons, or complex devices. The sequence 1-2-3 is repeated three times.

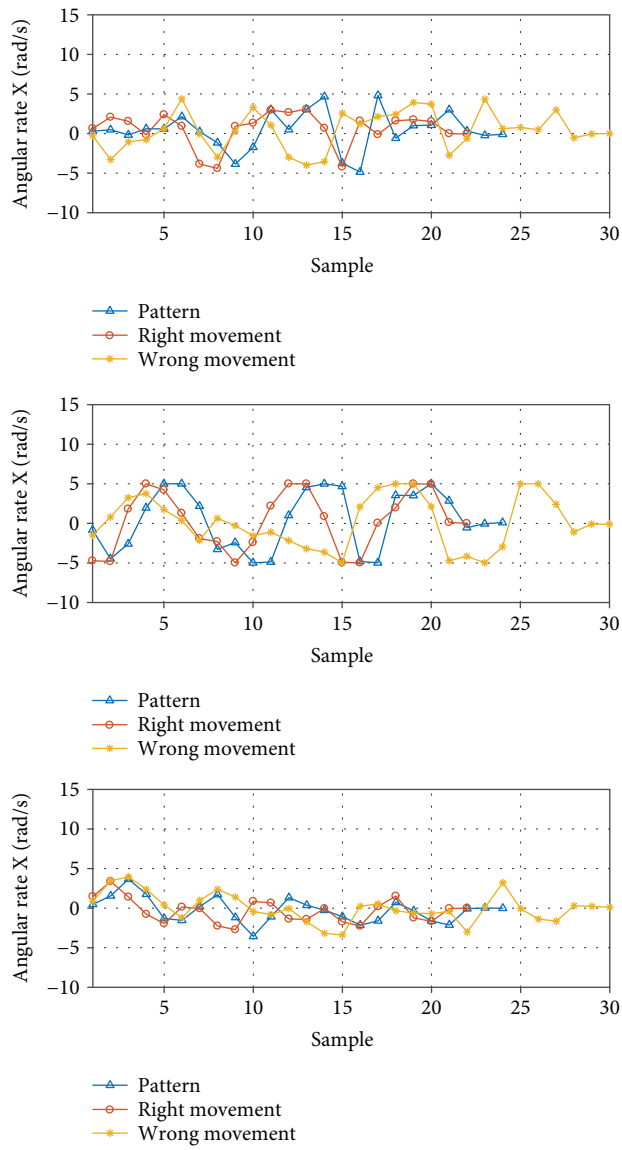


Figure 4: LoMoCA evaluation executing a right and a wrong movement compared to the pattern (gyroscope).

(v) *Notification*: enables operators to send an emergency message to notify their superior about any issue with no need to abandon the task at hand. It consists of moving the hand in vertical position (1-2 sequence) and shaking it two times (2-3-2 sequence repeated twice).

Assembly movement represents a typical production chain task, where different actions must be realized in the right order. The right sequence of actions is represented in Table 2, and to provide a perspective of data to be analysed by LoMoCA, we represented the angular rate samples of three movements overlapped (see Figure 4): *assembly pattern*, *assembly movement*, and *wrong assembly movement*. All movements have a duration of 2-3 seconds to highlight the low-frequency characterisation. *Assembly pattern* represents the pattern stored in memory to be used in the algorithm; a

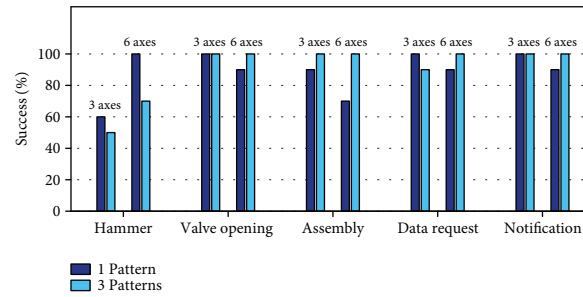


Figure 5: Number of pattern influence in LoMoCA.

right assembly movement follows the right sequence of actions, and *wrong assembly* refers to the execution of actions in a wrong sequence.

4.1.1. Number of Pattern Influence. LoMoCA introduces as many patterns as required to improve movement characterisation. Since there are many possible manners to execute a movement, the longer the movement is the lower similarity that is appreciated. The baseline idea consists of adding several patterns of the same movement.

The first experiment addresses the influence of the number of patterns stored in characterisation accuracy. Figure 5 shows a statistic with the success rate (using 3 or 6 DOF for each case). The experiment was based on 10 random sequences of the 5 movements recorded to obtain sufficient samples. The results revealed that using only accelerometers lead to wrong hammer characterisations which were identified as *data request* (see Figure 5). In that case, the use of 3 DOF (only acceleration) is not enough to distinguish between similar movements as *hammer* and *data request*. Thus, we concluded that a small number of axes combined with few sample movements increases similarity and consequently, makes LoMoCA select the shortest movement. Thus, the problem of using only acceleration data undermined accuracy increasing false negatives.

This issue was solved by adding the gyroscope data source (6 DOF using 3 linear and 3 rotational axes) to provide extra information, enabling LoMoCA to eradicate false negative characterisations. However, the number of patterns conditioned the success rate depending on the length of each movement (i.e., short movements like *hammer* do not require a large number of patterns to reach the 100% according to Figure 5). Therefore, the achievement of a high success rate to satisfy industry requirements is only possible by the use of gyroscope data in this case, being present the mentioned low-frequency restriction.

The addition of extra DOF by including gyroscopes is helpful for improving characterisation of simple movements, at the expense of a higher energy consumption. As stated in previous sections, the key priority of the system is reliability in order to enable the use of OperaBLE in Industry 4.0 scenarios. Thus, the extra power demand required by the addition of gyroscopes is justified to achieve a proper characterisation in the case of short movements. However, depending on the final application of OperaBLE, the use of 3 or 6 DOF is an aspect to be carefully

Table 3: Assembly successful recognition using 3 or 6 axes.

Sequence	3 axes		6 axes	
	Output	Success	Output	Success
1-2-3	✓	✓	✓	✓
(✓)	✓	✓	✓	✓
3-2-1	✓	✗	✗	✓
(✗)	✗	✓	✗	✓
2-3-1	✗	✓	✗	✓
(✗)	✗	✓	✗	✓
3-1-2	✗	✓	✗	✓
(✗)	✗	✓	✗	✓
2-1-3	✓	✗	✓	✗
(✗)	✓	✗	✗	✓
1-3-2	✓	✗	✗	✓
(✗)	✗	✓	✗	✓
		66.7%		91.7%

considered, according to the results of this experiment and the energetic restrictions.

4.1.2. Assembly Stress Test. The methodology of this experiment consisted of executing twice *assembly* in all possible wrong sequence combinations (3-2-1, 2-3-1, etc.), considering both 3 and 6 DOF to compare the influence of using more DOF in our algorithm. Considering the 1-2-3 sequence as correct (see Table 2), Table 3 provides the results of our *assembly* stress test where remarkable conclusions have been reached.

Using 3 DOF, the right movement (sequence 1-2-3) is correctly characterised in both repetitions. Nevertheless, we detected a false-positive issue since *assembly* was recognised in 4 occasions erroneously. Consequently, the addition of 3 extra DOF improved significantly the results, increasing the success rate from 66.7% to 91.7%. This fact enables the use of LoMoCA for more complex actions where the sequence of the movement is critical for the proper functioning of the process. This supports the idea that gyroscopes are essential, depending on the final application, to improve the performance of OperaBLE. Thus, only in situations where the addition of these 3 extra DOF is required to ensure reliability, the consumption of the device will be increased by the addition of a gyroscope data source.

4.1.3. Power Awareness: LoMoCA. This section offers an analysis of OperaBLE regarding power consumption during LoMoCA (varying number of DOF for different movements). The sampling frequency for consumption measurements was 10 Hz, and the source connected to OperaBLE was a 105 mAh battery. The experiment was performed during the execution of our reference movements (see Table 2); the numerical values were stored, and the consumption curve was obtained.

The average current measured during the characterisation stage for *hammer*, *valve opening*, and *assembly* movements was practically constant, since the duration of each movement had a greater impact on the consumption than

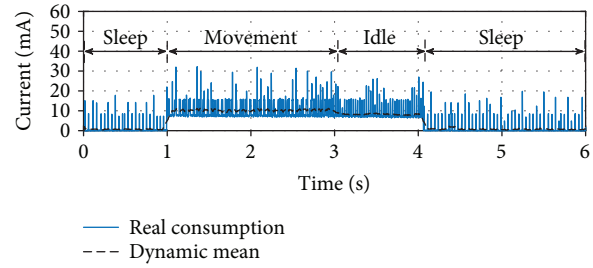


Figure 6: Average consumption per movement *(hammer)* in a 6-second interval using LoMoCA.

the number of iterations performed (determined by the kind of movement executed). Thus, our aim remained on the determination of the current consumed during the processing stage of a movement as well as the time interval required for characterisation.

The consumption curve of OperaBLE during *hammer* movement is shown in Figure 6, where the stages identified are *sleep*, *movement*, and *idle*. In this representation, OperaBLE characterised correctly the movement by only using 3 DOF (acceleration data). Additionally, it was maintained in a sleep state before and after characterisations (*movement* stage) to improve the identification of the different consumption stages, leaving a 1-second *idle* stage to ensure that the movement data gathering had finished. In the case of *valve opening* and *assembly* movements, the time spent during the movement reached approximately 3 seconds. Thus, OperaBLE remained 2 seconds in a *sleep* mode in these cases, where the sleeping time coincided with the movement time for a duty cycle.

The time spent during the movement characterisation is a relevant factor to consider. For the 6-second measurement interval defined in this experiment, OperaBLE consumes, at most (*assembly* movement), a capacity of 0.011 mAh. This means that our prototype will operate under such work flow for at least 15.70 hours while using 6 DOF (worst case). A lifespan of 10.27 hours is obtained dividing the 105 mAh capacity of our battery by the average consumption of 10.22 mA (movements being constantly characterised), which is the overall worst-case scenario. Even though this estimation is based on real measurements, the lifespan reached is ideal. For a 30% lifespan attenuation, the device would still operate for a complete working day with no energy constraints.

Apart from the time spent by each movement characterisation, this work proposes a set of priorities and decision-making procedure oriented to the number of DOF to be considered to save as much energy as possible. As stated in previous sections, the key priority for OperaBLE is reliability in order to fulfil the requirements of Industry 4.0. However, sustainability is as well an essential aspect for the algorithms, for which two main sources of energy consumption exist: the operating frequency of the microcontroller of the device (on which this work focuses) and the sensor modules attached to the device. Since communication management does not conflict with the previous, we highlight in the following lines the two

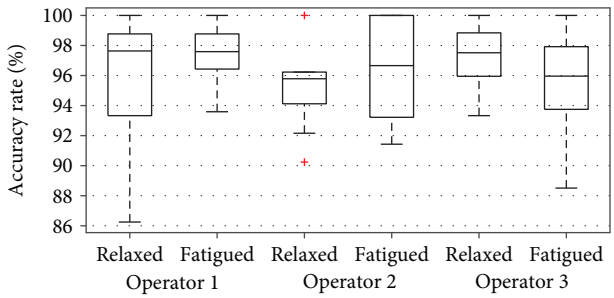


Figure 7: Statistical accuracy rate (%).

alternatives explored regarding the consumption sources for OperaBLE:

(1) The use of a high-frequency microcontroller, which would not require in many situations the addition of gyroscope data at the expense of a higher global consumption of the device. According to the technical datasheet of the IMU used, the consumption of the gyroscope is noticeably higher than the consumption of the accelerometer, which is the reason why avoiding the use of these extra DOF could contribute to greater energy savings in the sensor board. This alternative was discarded to focus on the microcontroller as the energy saving source for OperaBLE, which is the objective of this work and is expected to provide greater energy savings.

(2) The implementation of the system with a low-frequency limitation applied to the microcontroller, at the expense of adding in certain situations 3 extra DOF provided by the gyroscope to satisfy the reliability requirement. Although the sensor consumption would be increased, the global consumption of the microcontroller board can be considerably reduced, which affects the overall efficiency of the device. Focusing on this idea, this work proposes the reduction of power consumption based on the implementation of LoMoCA with this low-frequency restriction.

As a result, since the experiments carried out highlight a necessity of adding gyroscope data in short movements such as *hammer* (to satisfy the reliability requirement), the algorithm was developed to be easily adaptable to the final application. Thus, 3 DOF can be used in applications where LoMoCA does not require gyroscope data to successfully characterise movements (i.e., *valve opening*, *data request*, or *notification*) and 6 DOF will be used in situations where the reliability of characterisation may be undermined by this low-frequency requirement. On the whole, the sustainability of the proposal is based on the low-frequency operation of lightweight devices highly compatible with LoMoCA and enhanced by the addition or removal of angular velocity measurements depending on the necessities.

4.2. AHRA Experiments. This section focuses on AHRA, our adaptive heart rate algorithm, to test the accuracy of measurements provided by OperaBLE and the adaptability of the prototype to different workers and conditions.

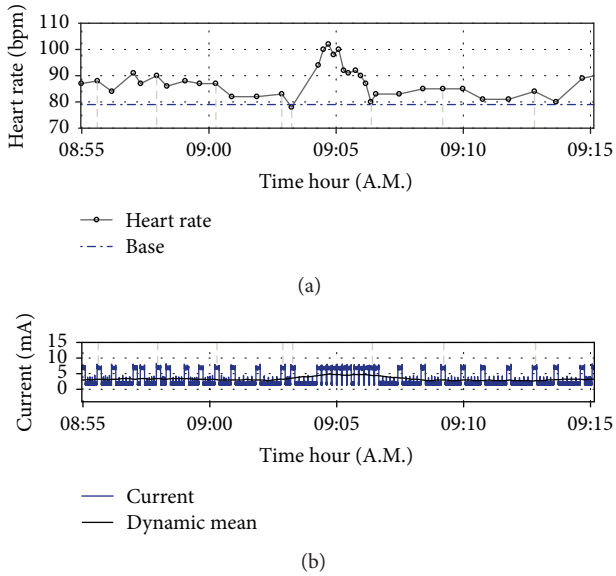


Figure 8: Heart rate test and simultaneous consumption.

4.2.1. Adaptive Heart Rate Algorithm Measurements. Three different workers were selected to carry out the first experiment, which consisted of extracting a bpm measurement while performing different tasks. The measurements where stored and compared to the ones calculated by characterising the average peak-to-peak time with a digital oscilloscope.

During the experiment, the calibration routine defined in AHRA was used with three different test workers, 10 times for each subject. Eventually, they were asked to fulfil a form indicating whether they had felt fatigued or relaxed. The percentage of statistical accuracy for each group of 10 measurements has been represented in Figure 7, where *fatigued* and *relaxed* categories are highlighted for each operator.

The success accuracy rate (see Figure 7) shows the proximity between the measured and the real bpm value. The results of the experiment have been classified into 6 groups according to the state of each operator. The highest dispersion is noticed in operator 1 *(relaxed)*, where the quartiles range from 86% to 100% of accuracy rate. The statistical representation of the *relaxed* state of operator 2 clearly shows that our algorithm has a good degree of repeatability (see quartiles 2 and 3 in Figure 7).

AHRA is suitable for measuring the heart rate in different conditions and workers. The measured value coincided with the expected in repeated occasions, and the average accuracy rate of success reached during the experiment was 96.16%. Since the main aim of OperaBLE is to identify risky situations, the algorithm is accurate and appropriate for detecting harmful changes to notify nearby supervisors.

4.2.2. Power Awareness: AHRA. This section describes an experiment based on a power-aware approach to quantify the energy savings of adapting sampling frequency to the

state of each worker. For this, a specific operator was subjected to differentiated emotional states triggered by external agents. The resulting data is shown in Figure 8(a), where a 20-minute time interval has been represented to focus on the most representative attribute of AHRA (adaptability). The part of Figure 8(b) includes the consumption obtained while measuring heart rate in OperaBLE to remark sustainability. It can be noticed how OperaBLE turns into sleep mode after each measurement, adjusting the sleeping interval (according to (1) and (2)).

Before starting to measure heart rate (08:54 A.M.), the operator was asked to calibrate OperaBLE to define a reference bpm value. The measurements shown were obtained from 08:55 A.M. to 09:15 A.M., time interval in which the test worker was performing a predefined task, and ten minutes after starting to store measurements (around 09:05 A.M.), an external agent provoked an unexpected situation. Consequently, the average bpm increased instantaneously and we could verify how the algorithm applied factor A (1) to reduce drastically the sleeping interval. Eventually, the subject returned to normal bpm values and OperaBLE augmented the sampling interval to save as much energy as possible (09:10 A.M.).

According to this experiment, we conclude that OperaBLE has an instantaneous current of 3 mA during relaxed situations while the consumption could be increased up to 7 mA when identifying fatigue or stress at work. Thus, OperaBLE achieves the sustainability challenge and ensures a minimum 15-hour lifespan (up to 35 hours in nonharmful situations).

5. Conclusions and Future Work

This research is motivated by the most challenging goals of the emerging industrial paradigm, the so-called Industry 4.0, where IoT technologies will play an essential role towards smart factories. In order to enhance working conditions and improve efficiency at work, this paper introduces our factory-wearable prototype, OperaBLE. This proposal is due to track operators in smart factories, avoiding accidents at workplaces and identifying harmful working conditions.

The major achievement of this work is the introduction of two power-aware algorithms, AHRA and LoMoCA, which have been specifically developed to operate in low-cost and low-power devices such as OperaBLE. These algorithms are based on flexible approaches in order to face the Industry 4.0 sustainability challenge. A set of experiments were carried out to evaluate these algorithms separately and verify the reliability and lifespan of the OperaBLE device during a working day.

Regarding LoMoCA, we reached the conclusion that the combination of accelerometers and gyroscopes (considering one pattern for simple movements and three for complex actions) conformed the best configuration achieving a 96% of success in an average situation, at the expense of greater energy consumption caused by gyroscopes. However, the use of this extra sensor is only justified in the case of specific situations such as short or sequence-sensitive movements, which could create confusion to the algorithm. Therefore, depending on the final application and the kind of movements to be characterised, a higher energy demand will be required not to compromise reliability for industrial applications, which is a key priority.

OperaBLE is flexible enough to characterise accurately any movement using low frequency, which enables its use in lightweight devices contributing towards a better use of the resources and an enhanced energy efficiency. The acceleration-based system developed permits data request operations by tapping on the surface of machines, avoiding the use of screens, or complex interfaces. It enables operators to send supervisors movement-based notifications immediately in case of emergency, which is essential to prevent accidents at work.

Regarding well-being in smart factory spaces, three workers were selected for evaluating AHRA in different states. Measurements presented an average accuracy rate of success 96.16% using OperaBLE with respect to measurements provided by a digital oscilloscope, which enabled OperaBLE to identify harmful conditions for workers such as stress or fatigue. Furthermore, thanks to the adaptable behaviour of AHRA, the average current consumed during heart rate measurements was minimal (3 mA approximately) while the low-frequency operation of LoMoCA allowed OperaBLE to reach a 10.27-hour lifespan in a worst-case scenario of constant movement characterisation and using a 105 mAh battery capacity.

Future works include a redesign of the prototype in order to reduce its size and enhance comfortability for operators. More algorithms are currently being developed to empower modularity in OperaBLE and track the surrounding environments of operators ensuring safety at work. Since BLE technology enriches scalability and offers possibilities for larger-scale applications, a BLE mesh network is being deployed at the moment for evaluating a complete BLE-based CPS. The results achieved are promising, deserving further research to evaluate robustness and increases productivity in real Industry 4.0 environments.

Data Availability

The data that support the findings of this study are available from the corresponding author, Diego Hortelano, upon reasonable request.

Conflicts of Interest

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

Acknowledgments

This work was supported by the Spanish Ministerio de Eduación, Cultura y Deportes under Grants TIN2015-66972-C5-2-R (MINECO/FEDER) and TIN2015-65845-C3-2-R and also by the I+D+i Universidad de Castilla-La Mancha plan with the European Social Fund.

References

[1] Y. Lu, "Industry 4.0: a survey on technologies, applications and open research issues," *Journal of Industrial Information Integration*, vol. 6, pp. 1–10, 2017.

[2] E. Borgia, "The Internet of Things vision: key features, applications and open issues," *Computer Communications*, vol. 54, pp. 1–31, 2014.

[3] R. Rajkumar, I. Lee, L. Sha, and J. Stankovic, "Cyber-physical systems: the next computing revolution," in *DAC '10 Proceedings of the 47th Design Automation Conference*, pp. 731–736, New York, NY, USA, June 2010.

[4] Z. A. Shhedi, A. Moldoveanu, F. Moldoveanu, and C. Taslitchi, "Real-time hand hygiene monitoring system for HAI prevention," in *2015 E-Health and Bioengineering Conference (EHB)*, pp. 1–4, Iasi, Romania, November 2015.

[5] F. Cai, C. Yi, S. Liu et al., "Ultrasensitive, passive and wearable sensors for monitoring human muscle motion and physiological signals," *Biosensors and Bioelectronics*, vol. 77, no. 15, pp. 907–913, 2016.

[6] K. Niazmand, K. Tonn, Y. Zhao et al., "Freezing of gait detection in Parkinson's disease using accelerometer based smart clothes," in *2011 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pp. 201–204, San Diego, CA, USA, November 2011.

[7] R. Brugarolas, J. M. Valero-Sarmiento, and A. Brna, "Wearable SpO_2 and sleep posture monitoring system for obstructive sleep apnea patients," in *2015 IEEE Virtual Conference on Applications of Commercial Sensors (VCACS)*, pp. 1–6, Raleigh, NC, USA, March 2015.

[8] R. Abbasi-kesbi and A. Nikfarjam, "A mini wearable wireless sensor for rehabilitation applications," in *2015 3rd RSI International Conference on Robotics and Mechatronics (ICROM)*, pp. 618–622, Tehran, Iran, October 2015.

[9] M. Nguyen, L. Fan, and C. Shahabi, "Activity recognition using wrist-worn sensors for human performance evaluation," in *2015 IEEE International Conference on Data Mining Workshop (ICDMW)*, pp. 164–169, Atlantic City, NJ, USA, November 2015.

[10] R. San-Segundo, J. Lorenzo-Trueba, B. Martinez-Gonzalez, and J. M. Pardo, "Segmenting human activities based on HMMs using smartphone inertial sensors," *Pervasive and Mobile Computing*, vol. 30, pp. 84–96, 2016.

[11] J. S. Park, S. Robinovitch, and W. S. Kim, "A wireless wristband accelerometer for monitoring of rubber band exercises," *IEEE Sensors Journal*, vol. 16, no. 5, pp. 1143–1150, 2016.

[12] K. J. Sheehan, B. R. Greene, C. Cunningham, L. Crosby, and R. A. Kenny, "Early identification of declining balance in higher functioning older adults, an inertial sensor based method," *Gait & Posture*, vol. 39, no. 4, pp. 1034–1039, 2014.

[13] R. C. Van Lummel, E. Ainsworth, U. Lindemann et al., "Automated approach for quantifying the repeated sit-to-stand using one body fixed sensor in young and older adults," *Gait & Posture*, vol. 38, no. 1, pp. 153–156, 2013.

[14] S. Urmat and M. E. Yalçın, "Design and implementation of an ARM based embedded system for pedestrian dead reckoning," in *2015 9th International Conference on Electrical and Electronics Engineering (ELECO)*, pp. 885–889, Bursa, Turkey, November 2015.

[15] W. Lee, E. Seto, K.-Y. Lin, and G. C. Migliaccio, "An evaluation of wearable sensors and their placements for analyzing construction worker's trunk posture in laboratory conditions," *Applied Ergonomics*, vol. 65, pp. 424–436, 2017.

[16] W. Lee, K.-Y. Lin, E. Seto, and G. C. Migliaccio, "Wearable sensors for monitoring on-duty and off-duty worker physiological status and activities in construction," *Automation in Construction*, vol. 83, pp. 341–353, 2017.

[17] J. Chen, J. Qiu, and C. Ahn, "Construction worker's awkward posture recognition through supervised motion tensor decomposition," *Automation in Construction*, vol. 77, pp. 67–81, 2017.

[18] R. Akhavian and A. H. Behzadan, "Smartphone-based construction workers' activity recognition and classification," *Automation in Construction*, vol. 71, pp. 198–209, 2016.

[19] M. Oliver, G. M. Schofield, H. M. Badland, and J. Shepherd, "Utility of accelerometer thresholds for classifying sitting in office workers," *Preventive Medicine*, vol. 51, no. 5, pp. 357–360, 2010.

[20] Punch Through, "LightBlue family2018, http://www.punchthrough.com/bean.

[21] K. Mikhaylov and J. Tervonen, "Optimization of microcontroller hardware parameters for wireless sensor network node power consumption and lifetime improvement," in *International Congress on Ultra Modern Telecommunications and Control Systems*, pp. 1150–1156, Moscow, Russia, October 2010.

[22] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 26, no. 1, pp. 43–49, 1978.

[23] C. F. S. Souza, C. E. P. Pantoja, and F. C. M. Souza, *Verificação de Assinaturas Offline Utilizando Dynamic Time Warping*, Universidade de Brasília, 2015.

[24] E. Langholz, "Dynamic time warping," 2014, http://github.com/langholz/dtw.

[25] D. Hortelano, T. Olivares, M. C. Ruiz, C. Garrido-Hidalgo, and V. López, "From sensor networks to Internet of Things. Bluetooth Low Energy, a standard for this evolution," *Sensors*, vol. 17, no. 2, pp. 1–31, 2017.

Hindawi
Journal of Sensors
Volume 2018, Article ID 6492592, 10 pages
https://doi.org/10.1155/2018/6492592

Research Article

Secure Data Encryption for Cloud-Based Human Care Services

Taehwan Park,[1] **Hwajeong Seo**,[2] **Sokjoon Lee**,[3] **and Howon Kim**[1]

[1]*School of Computer Science and Engineering, Pusan National University, San-30, Jangjeon-dong, Geumjeong-gu, Busan 609-735, Republic of Korea*
[2]*IT Engineering, Hansung University, 116 Samseong-Yoro-16-Gil, Seongbuk-gu, Seoul 136-792, Republic of Korea*
[3]*System Security Research Group, Electronics and Telecommunications Research Institute, Daejeon 34129, Republic of Korea*

Correspondence should be addressed to Howon Kim; howonkim@pusan.ac.kr

Received 22 February 2018; Revised 9 June 2018; Accepted 4 July 2018; Published 6 August 2018

Academic Editor: Mucheol Kim

Sensor network services utilize sensor data from low-end IoT devices of the types widely deployed over long distances. After the collection of sensor data, the data is delivered to the cloud server, which processes it to extract useful information. Given that the data may contain sensitive and private information, it should be encrypted and exchanged through the network to ensure integrity and confidentiality. Under these circumstances, a cloud server should provide high-speed data encryption without a loss of availability. In this paper, we propose efficient parallel implementations of Simeck family block ciphers on modern 64-bit Intel processors. In order to accelerate the performance, an adaptive encryption technique is also exploited for load balancing of the resulting big data. Finally, the proposed implementations achieved 3.5 cycles/byte and 4.6 cycles/byte for Simeck32/64 and Simeck64/128 encryption, respectively.

1. Introduction

At present, numerous human care services are used in hospitals and clinics according to the development of IoT (Internet of Things) technologies. To provide human care services in the United States, a service provider must follow the HIPAA (Health Insurance Portability and Accountability Act). The purpose of the HIPAA is to protect the privacy of medical information. For this reason, the U.S. Department of Health and Human Services (HHS) has stipulated that human care service providers must follow the HIPAA to enhance the protection of patients' private and health information. In the *HIPAA Compliance Guide* [1], the service provider must use data encryption on portable devices and computer networks during secure PHI (protected health information) transmission and storage processes. However, various types of devices, such as 8-bit, 16-bit, and 32-bit IoT devices as well as communication protocols such as Bluetooth and Wi-Fi, are used to provide human care services. These IoT devices for human care services operate in a resource-constrained environment. For this reason, existing block ciphers use a SPN (substitution and permutation network) architecture such as AES [2], which requires a considerable amount of memory to save S-boxes and the round keys and operations. To address this problem, numerous lightweight block ciphers use ARX (addition/AND, rotation, eXclusive-OR) operations such as SIMON, SPECK [3], or Simeck [4]. A Simeck family block cipher [4] is suitable for RFID (radio-frequency identification) sensor environments, though it has not been considered by any standardization institute to be included in a standard. However, the Simeck family block cipher [4] is lightweight according to ARX (AND, rotation, eXclusive-OR) operations and supports various block/key sizes. Moreover, it is suitable for RFID, sensor devices, and IoT end devices, all of which have resource-restricted environments. If Simeck family block ciphers are used in the end devices or sensors used in conjunction with various human care application services, cloud or service platform servers must deal with encrypted big data from various end devices or sensors in their human care application services as rapidly as possible to ensure the availability of their human care application services. At that time, the size of the big data can differ

Table 1: Simeck family block cipher specification.

Block cipher	Block size (bit)	Key size (bit)	Round (T)
Simeck32/64	32	64	32
Simeck48/96	48	96	36
Simeck64/128	64	128	44

according to the data transmission cycle of the end devices. To process big data on a cloud or service platform, the cloud or service platform servers must have high-performance modern 64-bit processors (high-frequency processors which are capable of supporting SIMD (single instruction multiple data)) such as an Intel Xeon processor or an i7 series processor for human care application services. To address these issues, in this paper we propose efficient parallel implementation methods of the Simeck family block cipher [4] using an Intel AVX2 (Advanced Vector Extension 2) SIMD and an efficient adaptive encryption method to enhance human care service availability based on Simeck family block cipher AVX2-optimized implementations which support various data block sizes.

The remainder of this paper is organized as follows. Section 2 discusses previous works related to the Simeck family block cipher, AVX2, and human care service security. We propose the efficient AVX2-optimized Simeck implementation and adaptive encryption method for big data encryption for use within a human care service in Section 3. Section 4 provides experimental and evaluation results from the proposed AVX2-optimized Simeck implementation and adaptive encryption approach. Section 5 provides the conclusion.

2. Related Works

In this section, we describe works related to Simeck family block cipher and Intel AVX2 SIMD and works related to human care service security.

2.1. Simeck Family Block Cipher. The Simeck family block cipher was proposed at CHES 2015 [4], and it succeeded the architecture of SIMON and SPECK [3] and therefore has similar encryption round functions with different numbers of rotation operations and bit rotation operations. The purpose of Simeck is suitability for use with a lightweight block cipher for hardware environments and RFID systems. There are three types of Simeck family block cipher, that is, Simeck32/64, Simeck48/96, and Simeck64/128, which have 32, 36, and 44 rounds, respectively. Table 1 presents the Simeck family block cipher specifications, including the block/key size (bit) and the number of rounds.

The Simeck family block cipher encryption and decryption round functions have ARX (AND (⊙), rotation (rotation left, (<<<)), eXclusive-OR (⊕)) operations. Figure 1 shows the encryption round function of the Simeck block cipher at the ith round. In Figure 1, l_i denotes the left word, r_i denotes the right word during the ith round, and k_i denotes

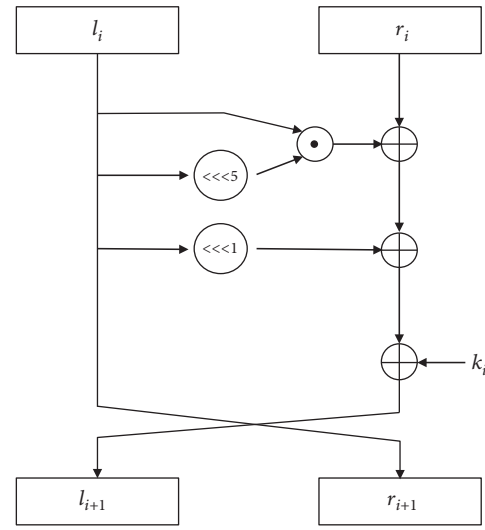


Figure 1: Simeck family block cipher encryption round function.

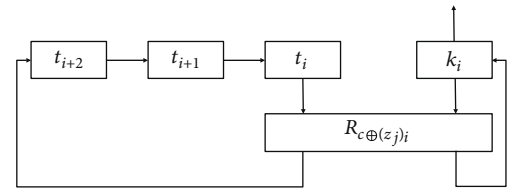


Figure 2: Simeck family block cipher key expansion.

the ith round key. The ith round function of the Simeck block cipher can be represented by the following equation:

$$R_{k_i}(l_i, r_i) = (r_i \oplus f(l_i) \oplus k_i, l_i),$$
$$f(x) = (x \odot \mathrm{ROL1}(x)) \oplus \mathrm{ROL5}(x). \quad (1)$$

In the equation of the Simeck encryption round function, the ROLr () function refers to the r-bit left rotation operation. It is expressed as <<<r in Figure 1. The Simeck encryption round function consists of 1 AND (⊙), 2 rotation left (<<<), and 3 eXclusive-OR (⊕) operations during each round.

In the Simeck family block cipher key schedule, Figure 2 shows the Simeck family block cipher key schedule as a block diagram. The key schedule of the Simeck family block cipher is similar to the key schedule procedure of the SIMON and SPECK family block ciphers [3]. $R_{C\oplus(z_j)_i}$ in Figure 2 is the Simeck round function with $C \oplus (z_j)_i$, which acts as the round key during each round. The round key k_i is generated from the master key K, which is initially segmented into four words and which loads as the initial states (t_2, t_1, t_0, k_0) of the feedback register architecture, as shown in Figure 2. The Simeck family block cipher key schedule uses initial states such as (1, 1, 1, 1, 1) and (1, 1, 1, 1, 1, 1). The first initial state is used with Simeck32/64 and Simeck48/96, and the second initial state is used with Simeck64/128. With regard to k_0, it is the least significant n bits of the master key K. To update the register values and generate a round key during each

round, it uses a round function with a round constant of $C \oplus (z_j)_i$ which is the bitwise eXclusive-OR result between the constant C and the i-bit of sequence Z_j, which can also act as the round key, that is, $R_{C\oplus(z_j)_i}$. The constant value (C) in the key schedule can be expressed as $2^n - 4 = 0 \times \mathrm{FF}, \ldots, \mathrm{FC}$. The $(z_j)_i$ in Figure 2 refers to the i-bit of sequence Z_j. There are two sequences: Z_0 and Z_1. Sequence Z_0 has 31 periods and can be generated using the primitive polynomial $x^5 + x^2 + 1$, while sequence Z_1 has 63 periods and can be generated using the primitive polynomial $x^6 + x + 1$. Sequence Z_0 is used with Simeck32/64 and Simeck48/96, and sequence Z_1 is used with Simeck64/128 for the Simeck family block cipher key schedule. The updating operation can be expressed as follows:

$$\begin{aligned} k_i &= t_i, \\ t_{i+3} &= k_i \oplus f(t_i) \oplus C \oplus (z_j)_i, \\ &\qquad 0 \le i \le T - 1. \end{aligned} \tag{2}$$

2.2. Related Works on Simeck Family Block Cipher. Related works which focus on the Simeck family block cipher can be divided into three types: cryptanalysis, side-channel attacks, and efficient implementations.

The first type involved cryptanalyses of the Simeck family block cipher. Kölbl and Roy [5] presented brief comparison results between SIMON and Simeck. Bagheri [6] proposed a linear cryptanalysis method and results with the reduced-round Simeck block cipher. Qiao et al. [7] proposed a differential analysis of Simeck using dynamic key-guessing techniques. Zhang et al. [8] proposed a zero-correlation linear cryptanalysis of Simeck. Wang [9] proposed a related-key differential analysis of the round-reduced Simeck block cipher based on mixed-integer linear programming. They found that a 14-round related-key differential distinguisher for Simeck32/64 is the best known method, and they used the dependencies of bitwise AND operations. Sadeghi and Bagheri [10] proposed an improved miss-in-the-middle approach to find zero-correlation linear distinguishers and impossible differentials on Simeck48 and Simeck64. They attacked 15-round Simeck48 and 17-round Simeck64 using their proposed zero-correlation linear approximation method which relies on the duality of the zero-correlation and impossible differential, and they attacked 27-round Simeck48 and 31-round Simeck64 based on their proposed zero-correlation linear distinguishers. Moreover, they proposed impossible differential attacks on 22-round Simeck48 and 24-round Simeck64 based on the impossible differential characteristics of Simeck.

The second category is side-channel attacks on the Simeck family block cipher. Qin et al. [11] proposed a linear hull attack with dynamic key-guessing techniques on round-reduced Simeck, similar to an approach by Qiao et al. [7]. Ryabko and Soskov [12] proposed a distinguishing attack on several lightweight block ciphers, including Simeck. Yoshikawa et al. [13] proposed a multiple-round-aware power analysis attack method and attack results on Simeck. Nozaki et al. [14] proposed an electromagnetic analysis of Simeck FPGA implementation using the linear relationship between the Hamming distance and electromagnetic waves, similar to DEMA and CEMA. Nalla et al. [15] proposed a differential fault attack on Simeck. They conducted random bit-flip fault attacks (requiring $n/2$ faults to recover the n-bit last round key) and a random byte fault attack (needing $n/6.5$ faults to recover the n-bit last round key) on Simeck. Nozaki et al. [16] proposed a double-round-driven electromagnetic analysis attack on Simeck FPGA implementation.

The final category contains methods related to the implementation of the Simeck block cipher. There are two efficient software implementations of Simeck on IoT-embedded devices: 8-bit AVR [17] and 16-bit MSP430 [18]. Simeck-based permutations for lightweight sponge cryptographic primitive hardware implementation were proposed at SAC 2017 [19].

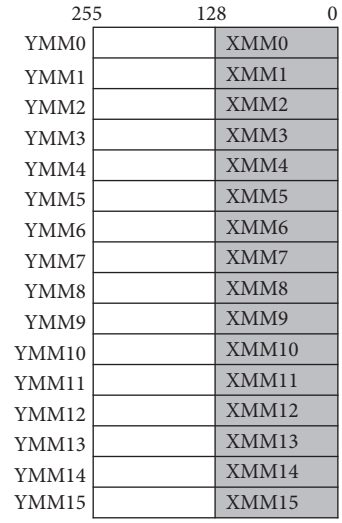


Figure 3: AVX2 registers.

2.3. Intel AVX2 SIMD. Intel AVX2 means Advanced Vector Extension 2 [20]. Intel has been providing AVX2 (Advanced Vector Extension 2) since the Intel Haswell architecture (Q2 2013). For AMD CPUs, the AMD Excavator processor (Q2 2015) and the Zen processor (Q1 2017) support AVX2. The main difference between AVX and AVX2 is that AVX2 supports 16 256-bit registers (YMM0–YMM16), as shown in Figure 3, with three-operand general-purpose bit manipulation and multiplication. Specifically, since the development of AVX2, vector shift operations are supported. If we want to use a rotation operation for an ARX-based block cipher, AVX2 shift operations can easily be used.

AVX2 registers of 256 bits can be written as 8-bit × 32, 16-bit × 16, and 32-bit × 8 for SIMD (single instruction multiple data), as shown in Figure 4. If we use a SIMD operation, we can calculate multiple instances of data simultaneously, implying that this approach can be used for multimedia processing. AVX2 SIMD supports intrinsic functions which correspond to AVX2 assembly instructions, and it is a type

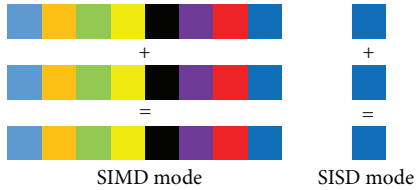


Figure 4: SIMD and SISD.

of API (application program interface). Accordingly, a developer can easily use Intel AVX2.

2.4. Related Works on Cryptographic Algorithm Implementation Based on AVX2. There are numerous research results pertaining to the implementation of cryptographic algorithms using AVX2.

Gueron and Krasnov [21] proposed multiprime RSA implementation methods and results using AVX2. In particular, they parallelized r modular exponentiations on RSA. Faz-Hernández and López [22] utilized efficient arithmetic operations on the prime field using AVX2, with performance benchmarked on the Intel Haswell processor. Faz-Hernández and López [23] proposed an efficient implementation of an elliptic curve (Curve25519) using AVX2. They proposed an accelerated prime field and elliptic curve arithmetic using AVX2. The Martins Paulo method [24] involves optimized fully homomorphic encryption implementation methods and results using AVX2. Cabral and López [25] proposed parallel SHA-3 family implementation methods and results using AVX2. They parallelized four digests from four different messages.

At present, AVX2 is used for implementing postquantum cryptography (PQC). Du et al. [26] proposed an efficient optimized number-theoretic transform (NTT) and high-precision discrete Gaussian sampler implementation methods and performance results for lattice-based public-key encryption using AVX2. Gueron and Schlieker [27] proposed optimized NTRUEncrypt implementation methods and results using AVX2. They also proposed replacing the SHA hash functions by pipelined AES-NI (Advanced Encryption Standard-New Instructions) for the rapid generation of randomness. Hamburg [28] proposed a new cryptosystem based on *integer module learning with errors* (I-MLWE), which uses the integers modulo ring with a generalized Mersenne prime number. In his paper, he proposed efficient software implementation methods and results using AVX2. Steinfeld et al. [29] proposed titanium, postquantum public-key encryption, and KEM. They proposed AVX2-optimized implementation methods and results with titanium. In PQCrypto 2017, fast lattice-based encryption SPRING implementation methods and results using AVX2 were posted. The security of lattice-based encryption SPRING relies on the hardness of the learning-with-rounding (LWR) problem. In these PQC implementation results using AVX2, the Hamburg method [28] and the approach developed by Steinfeld et al. [29] are focused on the NIST (National Institute of Standards and Technology) PQC standard competition. Accordingly, the efficient software implementation of a postquantum cryptographic algorithm using AVX2 is one of the important aspects of the NIST PQC standard competition.

Specifically, eBACs [30] offers SUPERCOP (System for Unified Performance Evaluation Related to Cryptographic Operations and Primitives), which measures the performance of hash functions, secret-key streams, public-key encryption, and other functions on modern processors using SSE, AVX, and AVX2.

2.5. Related Works on Human Care Service Security. There are two areas of related works on human care service security. The first is the security of commercial human care services, and the second area consists of research results on human care service security.

With regard to commercial human care services, Microsoft Azure, IBM Watson, and Amazon Web Services (AWS) are widely used. Microsoft published the "Microsoft Azure HIPAA/HITECH Act Implementation Guidance" publication [31]. MS Guidance follows the HIPPA/HITECH act and stores encryption keys separately. It also supports encryption-at-rest using .NET cryptographic services and SQL server encryption, including transparent data encryption with the Azure SQL database of PHI data. For communication security, SSL and TLS 1.1 are required. IBM Watson published the "Watson Developer Cloud Security Overview" [32]. In that document, they comply with the HIPPA Act and support authentication and authorization based on IBM Bluemix and end-to-end encryption following HTTPS (via TLS 1.2). Amazon Web Services (AWS) published "Architecting for HIPAA Security and Compliance on Amazon Web Services" [33]. They support PHI data encryption and protection in AWS during data transmission and storage using the AWS Key Management Service (KMS). They support SQL or Oracle database encryption for secure data storage and end-to-end encryption based on TLS or IPSec VPNs. AWS uses the AES-256 block cipher for data encryption on the database.

In research results on human care service security, Arunkumar and Anbuselvi [34] proposed secure cloud computing using the AES block cipher during PHI data storage in the cloud. Kumar et al. [35] proposed an IoT- and cloud-based patient monitoring system using the block ciphers of AES, DES, and Blowfish for data security. They suggested that the Blowfish block cipher is the most appropriate algorithm for their proposed health care system. Zhang et al. [36] described survey results on searchable encryption schemes and proposed searchable encryption to provide security and privacy on healthcare applications. Mohit et al. [37] proposed a mutual authentication protocol for a cloud-computing-based health care system. Mekala et al. [38] proposed a homomorphic encryption technique for healthcare multi-cloud computing which uses the Dynamo DB and Amazon Web Services (AWS). Mhatre et al. [39] compared attribute-based encryption for health records in cloud storage. They proposed multiauthority attribute-based encryption and provided an adequate, effective, and expressive solution to health record security problems. Zhao et al. [40] proposed attribute-based encryption with nonmonotonic access structures supporting fine-grained attribute

Table 2: AVX2 intrinsic functions for Simeck implementation.

Operations	AVX2 intrinsic functions
Load	*_mm256 loadu si256((m256i∗)x)*
Store	*_mm256 storeu si256((m256i∗)x,y)*
Set	*_mm256 set1 epi16(a)* *_mm256 set1 epi32(a)*
Bitwise AND	*_mm256 and si256(x,y)*
Bitwise OR	*_mm256 or si256(x,y)*
Bitwise XOR	*_mm256 xor si256(x,y)*
Shift left by r-bits	*_mm256 slli epi16(x,r)* *_mm256 slli epi32(x,r)*
Shift right by r-bits	*_mm256 srli epi16(x,r)* *_mm256 srli epi32(x,r)*
Rotation left by r-bits	*_mm256 or si256(_ mm256 slli epi16(x,r),* *_mm256 srli epi16(x,16-r))* *_mm256 or si256(_ mm256 slli epi32(x,r),* *_mm256 srli epi32(x,32-r))*

revocation in m-healthcare. However, they noted that their methods have problems such as large ciphertext sizes and lower efficiency rates.

3. Proposed Method

In this section, we propose efficient parallel implementation methods which use Simeck family block ciphers with AVX2 SIMD and adaptive encryption based on Simeck family block cipher AVX2-optimized implementation for human care services.

For the efficient parallel implementation of the Simeck family block cipher using AVX2, we used AVX2 intrinsic functions and optimized the AVX2 SIMD pipeline to avoid data hazards (aka: stall) and to enhance the performance level.

3.1. Efficient AVX2 Intrinsic Functions for Simeck. There are two methods for implementing the algorithm using AVX2. The first involves the use of AVX2 assembly instructions, and the second relies on AVX2 intrinsic functions. AVX2 intrinsic functions can correspond to AVX2 assembly instructions, as described earlier. We used the AVX2 intrinsic functions presented in Table 2. To implement the Simeck family block cipher, it is necessary to implement the data load from normal data to the AVX2 registers for the AVX2 SIMD operation, the data store from the AVX2 register to normal data, the data set to set the AVX2 register value as the Simeck encryption round key, and the bitwise AND, bitwise OR, bitwise XOR, shift left/right, and rotation operations for the Simeck encryption round function.

For the implementations of Simeck32/64 and Simeck 64/128, we used 256-bit AVX2 m256i data as follows: 16-bit × 16 and 32-bit × 8, respectively. AVX2 does not support vector rotation operations, and it only supports vector shift operations such as shift left/right via the r-bit intrinsic functions in Table 2. For this reason, it was necessary to create a rotation operation using the AVX2 vector shift left/right intrinsic functions and bitwise OR intrinsic functions shown in Table 2. For efficient rotation operations, we used 16-r and 32-r parts for left rotation operation by the r-bit AVX2 intrinsic functions shown in Table 2 as a precalculated constant value. For example, for rotation left by five bits on 16-bit × 16 AVX2 data x, we used the AVX2 intrinsic functions of *_mm256 or si256 (_mm256 slli epi16(x,5)* and *_mm256 srli epi16(x,11))*.

3.2. AVX2 SIMD Pipeline Optimization. For the optimization of the AVX2 SIMD pipeline, if we reused the AVX2 SIMD data which was the result data immediately before the operation as operand data during the next operation, it has a data dependency issue and therefore incurs the read-after-write (RAW) data hazard (aka: stall). This data hazard means that it requires several clock cycles to reload the data which was the result data immediately before the operation.

To avoid this data hazard, we rescheduled the order of operations for Simeck encryption. In the Simeck encryption round function shown in Figure 1, the order of operations is as follows: ROL5 (left) → AND (ROL5 (left), left) → XOR (right, AND (ROL5 (left), left)) → ROL1 (left) → XOR (ROL1 (left), XOR (right, AND (ROL5 (left), left))) → XOR (round key, XOR (right, AND (ROL5 (left), left))) → exchange left and right. If we implement the Simeck encryption round function in the above order, it can cause a data hazard. We rescheduled the order of operations as follows: ROL5 (left) → AND (ROL5 (left), left) → XOR (right, AND (ROL5 (left), left)) → ROL1 (left) → XOR (ROL1 (left), XOR (right, AND (ROL5 (left), left))) → round key Load → backup = left → XOR (round key, XOR (right, AND (ROL5 (left), left))) → right = backup to avoid data hazard.

For high performance during the encryption of big data, we maximize the usage of AVX2 registers by reusing AVX2 registers which are not used during each operation and considering AVX2 registers having plaintext blocks and round key blocks.

3.3. Adaptive Encryption for Human Care Service. Service servers for a human care service process massive amounts of data from various sensors or devices. However, the data sizes can differ in each case because each sensor or device has a different data transmission period. For this reason, it is necessary to encrypt data according to the data size each time when receiving or storing data in the database. To solve this problem, we proposed an adaptive encryption method for a human care service. Adaptive encryption can encrypt big data according to the data size. For adaptive encryption, we implemented AVX2-optimized Simeck32/64 and Simeck64/128 according to the encryption data block size. In the case of Simeck32/64, it can encrypt data from 16 blocks to 64 blocks. When it does this, each block size is 16 bits. Simeck64/128 can encrypt data from 8 blocks to 32 blocks. Each block size is 32 bits. Adaptive encryption uses the proposed AVX2-optimized Simeck32/64, and Simeck64/128 supports various block sizes for efficient data encryption of massive data according to the number of plaintext blocks.

Require: Plaintext Blocks $P \in \{P_1, P_2 \cdots, P_{blkNum}\}$ Round keys *RK*, The Number of Plaintext Blocks *blkNum*, Encryption Type *t*
Ensure: Ciphertext Blocks $C \in \{C_1, C_2, \cdots C_{blkNum}\}$
1: if t == 1 **then**
2: let *n1* =*blkNum*/64; and *r1* =*blkNum* %64;
3: let *n2* =*r1*/48; and *r2* =*r1*%48;
4: let *n3* =*r2*/32; and *r3* =*r2*%32;
5: let *n4* =*r3*/16; and *r4* =*r3*%16;
6: **if** r4 ≥1 **then**
7: *n4* ++;
8: *Pad* ();
9: **for** *i* from 1 to *n1* **do**
10: *Simeck* 32/64 *Enc SIMD* 64*Blks* $(RK, \_P_{partial_i}, C_{partial_i})$;
11: **for** *i* from 1 to *n2* **do**
12: *Simeck* 32/64 *Enc SIMD* 48*Blks* $(RK, \_P_{partial_{n1+i}}, C_{partial_{n1+i}})$;
13: **for** *i* from 1 to *n3* **do**
14: *Simeck* 32/64 *Enc SIMD* 32*Blks* $(RK, \_P_{partial_{n2+i}}, C_{partial_{n2+i}})$;
15: **for** *i* from 1 to *n4* **do**
16: *Simeck* 32/64 *Enc SIMD* 16*Blks* $(RK, \_P_{partial_{n3+i}}, C_{partial_{n3+i}})$;
17: **else**
18: let *n1* =*blkNum*/32; and *r1* =*blkNum* %32;
19: let *n2* =*r1*/24; and *r2* =*r1*%24;
20: let *n3* =*r2*/16; and *r3* =*r2*%16;
21: let *n4* =*r3*/8; and *r4* =*r3*%8;
22: **if** r4 ≥1 **then**
23: *n4* ++;
24: *Pad* ();
25: **for** *i* from 1 to *n1* **do**
26: *Simeck* 64/128 *Enc SIMD* 32*Blks* $(RK, \_P_{partial_i}, C_{partial_i})$;
27: **for** *i* from 1 to *n2* **do**
28: *Simeck* 64/128 *Enc SIMD* 24*Blks* $(RK, \_P_{partial_{n1+i}}, C_{partial_{n1+i}})$;
29: **for** *i* from 1 to *n3* **do**
30: *Simeck* 64/128 *Enc SIMD* 16*Blks* $(RK, \_P_{partial_{n2+i}}, C_{partial_{n2+i}})$;
31: **for** *i* from 1 to *n4* **do**
32: *Simeck* 64/128 *Enc SIMD* 8*Blks* $(RK, \_P_{partial_{n3+i}}, C_{partial_{n3+i}})$;
33 *Return C*

Algorithm 1: Adaptive encryption based on Simeck.

Algorithm 1 describes the proposed adaptive encryption algorithm. Input data for the adaptive encryption algorithm are plaintext blocks $P \in \{P_1, P_2, \ldots, P_{\text{blkNum}}\}$, the Simeck encryption round key RK, the number of plaintext blocks blkNum, and the encryption type *t*. The encryption type *t* indicates which Simeck family block cipher will be used for adaptive encryption. If it has a value of 1, this indicates that Simeck32/64 encryption will be used for adaptive encryption. The number of plaintext blocks blkNum refers to the total number of plaintext blocks, and plaintext blocks consist of blkNum blocks. Adaptive encryption using Simeck32/64 or Simeck64/128 has a similar algorithm routine. If the user wants to use Simeck32/64, it supports four types of blocks: 64, 48, 32, and 16 bits. Therefore, during the adaptive encryption process, this method determines how many encryption operations to run according to the block number and the number of plaintext blocks. From line 2 to line 5, it calculates *n*1, *n*2, *n*3, and *n*4 which refer to the number of encryptions for each supported block number. In the case of *n*4, if the remainder value *r*4 is 1 or more, it means that there are a number of plaintext blocks for encryption (less than 16 blocks), and we increase the value of *n*4 and use data padding to adjust the number of blocks used in the encryption process. In the case of data padding, there are many padding standards, such as PKSC7 and X923, among others. We assume that the padding method follows the relevant standard. After calculating the number of encryptions at each supported block number, we encrypt plaintext blocks according to the calculated number using each AVX2-optimized Simeck encryption supported fixed block size.

4. Experiment and Evaluation

In this section, we describe the experimental environment, procedures, and analysis method used to assess the performance of the proposed AVX2-optimized Simeck family block cipher implementation and adaptive encryption techniques.

Table 3: Proposed Simeck family block cipher performance.

Cipher	Number of blocks	Cycles/byte
Simeck32/64	16	5.3125
	32	3.9063
	48	3.5417
	64	3.5859
Simeck64/128	8	6.8750
	16	4.6875
	24	4.6146
	32	4.6875

4.1. Experiment. We developed the proposed AVX2-optimized Simeck family block cipher using AVX2 intrinsicfunctions and C language. We used GCC compiler version 5.4.0 with the compile options of *-O3-fomit-frame-pointer-mavx2-march = nativ-std = c99-mtune = native-fwrapv-funroll-loops* to optimize the compiling process. We conducted the experiment on a computer with the following specifications: Ubuntu 16.04.3 LTS 64 bit and Intel (R) Core (TM) i7-6700 CPU (@3.40 GHz, 32 GB RAM). We measured the average performance of 10,000,000 times encryption to ensure an accurate performance measurement.

4.2. Evaluation of AVX2-Optimized Simeck Implementation Performance. Table 3 describes the performance of the proposed AVX2-optimized Simeck family block cipher implementation. The performance is calculated in units of cycles/byte.

As shown in Table 3, Simeck32/64 encryption for 48 blocks and Simeck64/128 encryption for 34 blocks have the best performance, with 3.5417 cycles/byte and 4.6146 cycles/byte, respectively. Simeck32/64 encryption for 16 blocks and Simeck64/128 encryption for 8 blocks show lower performance outcomes because these encryptions encrypt only one 256-bit AVX2 register value. Hence, the performance in these cases is slightly lower. In the cases of Simeck32/64 encryption for 64 blocks and Simeck64/128 encryption for 32 blocks, these encryptions result in lower performance than Simeck32/64 encryption for 48 blocks and Simeck64/128 encryption for 32 blocks. Although Simeck32/64 encryption for 64 blocks and Simeck64/128 encryption for 32 blocks use the maximum number of AVX2 registers for efficiency, during the compile procedure, the GCC compiler does not use the AVX2 register designated by the C source code (AVX2 intrinsic functions).

For an objective evaluation of the Simeck AVX2-optimized implementation with the proposed methods, we implemented the SIMON family block cipher [3] with the proposed methods and measured the performance of the SIMON family block cipher with the proposed methods on the same environment. Table 4 describes the performances of the SIMON family block cipher with the proposed methods. SIMON32/64 encryption on 64 blocks has 4.1797 cycles/byte but the proposed Simeck32/64 encryption method on 64 blocks runs at 3.5859 cycles/byte (the best performance of Simeck32/64 is 3.5417 cycles/byte on 48 blocks of encryption) while SIMON64/128 encryption on 32 blocks operates at 5.7031 cycles/byte but proposed Simeck64/128 encryption on the same block size has 4.6875 cycles/byte (the best performance of Simeck64/128 encryption on 24 blocks runs at 4.6146 cycles/byte). The cause of the performance difference is the difference between the number of rotations left by the r-bit operations for Simeck and SIMON encryption round functions. The SIMON encryption round function is $R_k(x, y) = (y \oplus f(x) \oplus k, x)$, $f(x) = (Sx \& S^8 x) \oplus S^2 x$, and it requires three rotation left operations by r-bit ($S^r x$, r can be 1 bit, 8 bits, or 2 bits), three bitwise eXclusive-OR operations ($\oplus$), and one bitwise AND operation (&). However, the Simeck encryption round function is $R_k(x, y) = (y \oplus f(x) \oplus k, x)$, $f(x) = (x \& S^5 x) \oplus S^1 x$, and it requires two rotation left operations by r-bit (r can be 1 bit or 5 bits), three bitwise eXclusive-OR operations, and one bitwise AND operation. For this reason, the proposed Simeck32/64 AVX2-optimized implementations show performance improvements of 23.53%, 25.20%, 26.47%, and 16.56% over SIMON32/64 with the proposed methods for each respective number of blocks. With the proposed Simeck64/128 AVX2-optimized implementations, the performance improvements are 28.64%, 43.33%, 30.02%, and 21.67% over SIMON64/128 with the proposed methods for each corresponding number of blocks. Specifically, if comparing the best performance between the proposed Simeck and SIMON with the proposed methods, the best performance of the proposed Simeck32/64 and Simeck64/128 shows performance improvements of 18.01% and 23.59%, respectively, over the best performance of SIMON32/64 and SIMON64/128 with the proposed methods.

Table 4: SIMON family block cipher with proposed methods performance.

Cipher	Number of blocks	Cycles/byte
SIMON32/64	16	6.5625
	32	4.8906
	48	4.4792
	64	4.1797
SIMON64/128	8	8.8438
	16	6.7188
	24	6.0000
	32	5.7031

Table 5: Performance comparison result.

Cipher	Cycles/byte
SIMON64/128 for 64 bytes [30]	7.72
Proposed Simeck64/128 for 16 blocks (64 bytes)	*4.6875*

We also compared the performance capabilities between the proposed AVX2-optimized Simeck64/128 implementation for 16 blocks (64 bytes) and the SIMON64/128 for 64 bytes on eBACS SUPERCOP [30] (https://bench.cr.yp.to/results-stream.html). The SIMON64/128 for 64 bytes on

Table 6: Predicted performance of adaptive encryption.

Type of adaptive Enc.	Cycles/byte
Adaptive Enc.(Simeck32/64)	$(3.5859 \times n1) + (3.5417 \times n2) + (3.9063 \times n3) + (5.3125 \times n4)$
Adaptive Enc.(Simeck64/128)	$(4.6875 \times n1) + (4.6146 \times n2) + (4.6875 \times n3) + (6.8750 \times n4)$

Table 7: Comparison evaluation between previous works and the proposed method.

Method	Encryption	Optimized implementation	Optimized for big data
Arunkumar et al. [34]	AES	X	X
Kumar et al. [35]	AES, DES, Blowfish	X	X
Zhang et al. [36]	Searchable encryption	X	X
Zhao et al. [40]	Attribute-based encryption	X	X
Proposed method	*Simeck*	*O*	*O*

eBACS SUPERCOP [30] is implemented using AVX2. The performance of the SIMON64/128 for 64 bytes on eBACS SUPERCOP [30] is measured on systems using the AMD64, Kaby Lake (906e9), and 2017 Intel Xeon E3-1220 v6 processors with 4× in a 3000 MHz environment. Table 5 describes the performance comparison results between the proposed AVX2-optimized Simeck64/128 implementation for 16 blocks (64 bytes) and SIMON64/128 for 64 bytes on eBACS SUPERCOP [30]. The proposed AVX2-optimized Simeck64/128 implementation for 16 blocks (64bytes) is approximately 60.72% faster than the SIMON64/128 for 64 bytes on eBACS SUPERCOP [30] with the same data size (64 bytes).

4.3. Evaluation of Adaptive Encryption. Table 6 describes the predicted performance of the proposed adaptive encryption approach based on the performance of the AVX2-optimized Simeck family block cipher implementation. If the total number of plaintext blocks is N, it can be written as $N = n1 + n2 + n3 + n4$. Here, $n1$ refers to the number of Simeck32/64 encryptions for 64 blocks or Simeck64/128 for 32 blocks, $n2$ is the number of Simeck32/64 encryptions for 48 blocks or Simeck64/128 encryption for 24 blocks, $n3$ is the number of Simeck32/64 encryptions for 32 blocks or Simeck64/128 encryptions for 16 blocks, and $n4$ denotes the number of Simeck32/64 encryptions for 16 blocks or Simeck64/128 encryptions for eight blocks according to Algorithm 1. We can then predict the performance of the proposed adaptive encryption by aggregating the product of each number of Simeck encryptions for specific blocks and each proposed AVX2-optimized Simeck encryption performance outcome, as shown in Table 6.

The advantage of the proposed adaptive encryption becomes apparent when the number of plaintext blocks is increased. In such a case, the performance is also increased. Therefore, the proposed adaptive encryption method with the proposed AVX2-optimized Simeck family block cipher implementation is an adequate and efficient data encryption method which can be used by human care services to comply with the HIPPA security requirement (data encryption).

4.4. Comparison Evaluation. We compared previous works on human care service security and the proposed method from optimized implementation and optimized for big data perspectives. Table 7 describes the comparison evaluation results. From the perspective of optimized implementation, previous works focused on applying an encryption method to a human care service for security. Accordingly, they did not optimize the encryption implementation step. However, with the proposed methods, these are optimized using AVX2 SIMD considering the SIMD implementation condition, and the Simeck family block cipher outperforms AES [4]. From the perspective on optimization for big data, previous works such as those by Arunkumar et al. [34] and Kumar et al. [35] applied encryption to a cloud computing environment without considering the processing of big data. However, the proposed adaptive encryption method considers efficient big data processing on a cloud environment or on the server side to provide security (data encryption) with good availability following the HIPAA security requirements for a human care service.

5. Conclusion

In a human care service, service providers must comply with the HIPAA security and privacy requirement in the United States for security. For this reason, human care service providers such as MS Azure, IBM Watson, and Amazon Web Service provide security following HIPAA, and there are many research results on encryption for a human care service. However, previous research results focused generally on applying encryption methods to a human care service without considering efficiency and availability. If lightweight block ciphers are used at the end devices or sensors for various human care application services, cloud or service platform servers must deal with encrypted big data from the various end devices or sensors of their human care application service as rapidly as possible to ensure the availability of their human care application service considering different received data sizes according to the data transmission cycles of the end devices. In this paper, to solve these problems, we proposed AVX2-optimized Simeck family block cipher implementations supporting various numbers of blocks with good performance to provide efficiency and availability of data encryption at the cloud or on the server side. The

proposed AVX2-optimized Simeck32/64 encryption for 48 blocks has 3.5417 cycles/byte while Simeck64/128 encryption for 24 blocks has 4.6146 cycles/byte. For an objective evaluation of the proposed methods, we compared Simeck family block cipher AVX2-optimized encryption and SIMON family block cipher AVX2-optimized encryption based on the proposed method. The best performance outcomes with the proposed Simeck32/64 and Simeck64/128 were correspondingly 18.01% and 23.59% performance improvements as compared to the best performance of SIMON32/64 and SIMON64/128 with the proposed methods. Specifically, the proposed AVX2-optimized Simeck64/128 implementation method for 16 blocks (64 bytes) is approximately 60.72% faster than the SIMON64/128 for 64 bytes on eBACS SUPERCOP [30] with the same data size (64 bytes). We also proposed adaptive encryption based on AVX2-optimized Simeck encryption for efficient big data encryption on the cloud or server side to ensure high performance. The strong points of the proposed methods are AVX2-optimized Simeck block cipher implementation and an efficient adaptive encryption method for efficient big data encryption to enhance the availability of a human care application service. However, a limitation of the proposed methods is that it has not been applied to an actual human care service with optimized Simeck implementation using AVX2. To address this issue, we will apply the proposed methods on IBM Bluemix and IBM Watson for a human care service and conduct research on the efficient AVX-512-optimized implementation of the Simeck family block cipher in the future.

Data Availability

The proposed Simeck family block cipher implementation source codes are uploaded to the GitHub repository (https://github.com/pth5804/Simeck_AVX2). SIMON family block cipher implementation source codes based on the proposed methods are also uploaded to the GitHub repository (https://github.com/pth5804/SIMON_AVX2).

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work of Taehwan Park and Howon Kim was supported by the Institute for Information & Communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (no. 2017-0-01791, Development of Security Technology for Energy Platform and Device). This work of Hwajeong Seo was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (no. NRF-2017R1C1B5075742). This work of Sokjoon Lee was supported by Institute for Information & Communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (no. B0717-16-0097, Development of V2X Service Integrated Security Technology for Autonomous Driving Vehicle). This paper was proofread by the KAIST Language Center.

References

[1] K. Andrew, A. Steve, and G. Shane, *HIPAA Compliance Guide*, Tech. Rep. HIPAA Compliancy Group, 2017.

[2] J. Daemen and V. Rijmen, *The Design of Rijndael: AES-The Advanced Encryption Standard*, Springer Science & Business Media, 2013.

[3] R. Beaulieu, S. Treatman-Clark, D. Shors, B. Weeks, J. Smith, and L. Wingers, "The SIMON and SPECK lightweight block ciphers," in *2015 52nd ACM/EDAC/IEEE Design Automation Conference (DAC)*, pp. 1–6, San Francisco, CA, USA, June 2015.

[4] G. Yang, B. Zhu, V. Suder, M. D. Aagaard, and G. Gong, "The Simeck family of lightweight block ciphers," in *International Workshop on Cryptographic Hardware and Embedded Systems*, T. Güneysu and H. Handschuh, Eds., Springer, Berlin, Heidelberg, 2015.

[5] S. Kölbl and A. Roy, "A brief comparison of Simon and Simeck," in *International Workshop on Lightweight Cryptography for Security and Privacy*, Springer, Cham, 2016.

[6] N. Bagheri, "Linear cryptanalysis of reduced-round Simeck variants," in *International Conference in Cryptology in India*, Springer, Cham, 2015.

[7] K. Qiao, L. Hu, and S. Sun, "Differential analysis on Simeck and Simon with dynamic key-guessing techniques," in *International Conference on Information Systems Security and Privacy*, Springer, Cham, 2016.

[8] K. Zhang, J. Guan, B. Hu, and D. Lin, "Security evaluation on Simeck against zero-correlation linear cryptanalysis," *IET Information Security*, vol. 12, no. 1, pp. 87–93, 2018.

[9] S. Wang, "Related-key differential analysis of round-reduced Simeck," in *2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, Palladam, India, February 2017.

[10] S. Sadeghi and N. Bagheri, "Improved zero-correlation and impossible differential cryptanalysis of reduced-round Simeck block cipher," *IET Information Security*, vol. 12, no. 4, pp. 314–325, 2018.

[11] L. Qin, H. Chen, and X. Wang, "Linear hull attack on round-reduced Simeck with dynamic key-guessing techniques," in *Australasian Conference on Information Security and Privacy*, Springer, Cham, 2016.

[12] B. Ryabko and A. Soskov, *The Distinguishing Attack on Speck, Simon, Simeck, HIGHT and LEA*, Cryptology ePrint, 2018.

[13] M. Yoshikawa, Y. Nozaki, and K. Asahi, "Multiple rounds aware power analysis attack for a lightweight cipher Simeck," in *2016 IEEE Second International Conference on Big Data Computing Service and Applications (Big Data Service)*, Oxford, UK, 2016.

[14] Y. Nozaki, Y. Ikezaki, and M. Yoshikawa, "Tamper resistance of IoT devices against electromagnetic analysis," in *2016 IEEE International Meeting for Future of Electron Devices, Kansai (IMFEDK)*, Kyoto, Japan, June 2016.

[15] V. Nalla, R. A. Sahu, and V. Saraswat, "Differential fault attack on Simeck," in *Proceedings of the Third Workshop on Cryptography and Security in Computing Systems-CS2'16*, Prague, Czech Republic, January 2016.

[16] Y. Nozaki, Y. Ikezaki, and M. Yoshikawa, "Double-rounds-driven electromagnetic analysis attack for a lightweight block cipher Simeck and its evaluation," *Electronics and Communications in Japan*, vol. 100, no. 12, pp. 29–38, 2017.

[17] T. Park, H. Seo, B. Bae, and H. Kim, "Efficient implementation of Simeck family block cipher on 8-bit processor," *Journal of Information and Communication Convergence Engineering*, vol. 14, no. 3, pp. 177–183, 2016.

[18] T. Park, H. Seo, G. Lee, and H. Kim, "Efficient implementation of Simeck family block cipher on 16-bit MSP430," in *2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN)*, Milan, Italy, July 2017.

[19] R. AlTawy, R. Rohit, M. He, K. Mandal, G. Yang, and G. Gong, "sLiSCP: Simeck-based permutations for lightweight sponge cryptographic primitives," in *International Conference on Selected Areas in Cryptography*, pp. 129–150, Springer, Cham, 2017.

[20] G. Lento, *Optimizing Performance with Intel Advanced Vector Extensions*, 2014, online.

[21] S. Gueron and V. Krasnov, "Speed records for multi-prime RSA using AVX2 architectures," in *Information Technology: New Generations*, pp. 237–245, Springer, Cham, 2016.

[22] A. Faz-Hernández and J. López, "On software implementation of arithmetic operations on prime fields using AVX2," *XIV Simpósio Brasileiro em Segurança da Informaçao e de Sistemas Computacionais*, pp. 338–341, 2014.

[23] A. Faz-Hernández and J. López, "Fast implementation of Curve 25519 using AVX2," in *International Conference on Cryptology and Information Security in Latin America*, Springer, Cham, 2015.

[24] P. Martins and L. Sousa, "Enhancing data parallelism of fully homomorphic encryption," in *International Conference on Information Security and Cryptology*, Springer, Cham, 2016.

[25] R. Cabral and J. López, "Software implementation of SHA-3 family using AVX2," *Simpósio Brasileiro em Segurança da Informaçao e de Sistemas Computacionais*, pp. 330–333, 2014.

[26] C. Du, G. Bai, and H. Chen, "Towards efficient implementation of lattice-based public-key encryption on modern CPUs," in *2015 IEEE Trustcom/BigDataSE/ISPA*, Helsinki, Finland, August 2015.

[27] S. Gueron and F. Schlieker, "Software optimizations of NTRUEncrypt for modern processor architectures," in *Information Technology: New Generations*, pp. 189–199, Springer, Cham, 2016.

[28] M. Hamburg, *Integer Module LWE Key Exchange and Encryption: the Three Bears*, Rambus, Inc., 2017.

[29] R. Steinfeld, A. Sakzad, and R. K. Zhao, *Titanium: Proposal for a NIST Post-Quantum Public-Key Encryption and KEM Standard*, Monash University, 2017.

[30] D. J. Bernstein and T. Lange, *eBACS: ECRYPT Benchmarking of Cryptographic Systems*February 2018, https://bench.cr.yp.to.

[31] Microsoft, "Microsoft Azure HIPAA/HITECH Act implementation guidance," Tech. Rep., Microsoft, 2015.

[32] IBM, "Watson Developer Cloud security overview," Tech. Rep., IBM, 2016.

[33] A. W. Services, "Architecting for HIPAA security and compliance on Amazon Web Services," Tech. Rep., Amazon Web Services, 2018.

[34] R. J. Arunkumar and R. Anbuselvi, "Enhancement of cloud computing security in the health care sector," *International Journal of Computer Science and Mobile Computing*, vol. 6, no. 8, pp. 23–31, 2017.

[35] B. V. Kumar, M. Ramaswami, and P. Swathika, "Data security on patient monitoring for future healthcare application," *International Journal of Computer Applications*, vol. 163, no. 6, pp. 20–23, 2017.

[36] R. Zhang, R. Xue, and L. Liu, "Searchable encryption for healthcare clouds: a survey," *IEEE Transactions on Services Computing*, p. 1, 2017.

[37] P. Mohit, R. Amin, A. Karati, G. P. Biswas, and M. K. Khan, "A standard mutual authentication protocol for cloud computing based health care system," *Journal of Medical Systems*, vol. 41, no. 4, p. 50, 2017.

[38] M. Bhaskar, M. Ashok, and M. D. A. Hasan, "Homomorphic encryption algorithm used in multi-format data in collaborated healthcare multi cloud computing," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, vol. 2, no. 6, pp. 80–84, 2017.

[39] S. Mhatre, A. V. Nimkar, and S. N. Dhage, "Comparative study on attribute-based encryption for health records in cloud storage," in *2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, Bangalore, India, May 2017.

[40] Y. Zhao, P. Fan, H. Cai, Z. Qin, and H. Xiong, "Attribute-based encryption with non-monotonic access structures supporting fine-grained attribute revocation in M-healthcare," *International Journal of Network Security*, vol. 19, no. 6, pp. 1044–1052, 2017.

Hindawi
Journal of Sensors
Volume 2018, Article ID 6161978, 14 pages
https://doi.org/10.1155/2018/6161978

Research Article

Human Emotional Care Purposed Automatic Remote Portrait Drawing Generation and Display System Using Wearable Heart Rate Sensor and Smartphone Camera with Depth Perception

Gyeoungrok Lee,[1] **Dongwann Kang**,[2] **and Kyunghyun Yoon**[1]

[1]*School of Computer Science and Engineering, Chung-Ang University, Seoul, Republic of Korea*
[2]*Faculty of Science and Technology, Bournemouth University, Poole, UK*

Correspondence should be addressed to Kyunghyun Yoon; khyoon@cglab.cau.ac.kr

Received 23 February 2018; Revised 30 May 2018; Accepted 13 June 2018; Published 2 August 2018

Academic Editor: Ka L. Man

We propose a system that automatically generates portrait drawings for the purpose of human emotional care. Our system comprises two parts: a smartphone application and a server. The smartphone application enables the user to take photographs throughout the day while acquiring heart rates from the smartwatch worn by the user. The server collects the photographs and heart rates and displays portrait drawings automatically stylized from the photograph for the most exciting moment of the day. In the system, the user can recall the exciting and happy moment of the day through admiring the drawings and heal the emotion accordingly. To stylize photographs as portrait drawings, we employ nonphotorealistic rendering (NPR) methods, including a portrait etude stylization proposed in this paper. Finally, the effectiveness of our system is demonstrated through user studies.

1. Introduction

Internet of Things (IoT) has undergone rapid advances following the popularization of smartphones and wearable sensor devices. Owing to these circumstances, IoT has become nearly ubiquitous in our lives, widely supporting facets of our well-being as well. Especially, in healthcare front, many applications (e.g., Apple's HealthKit) based on IoT have been developed and employed. However, such applications generally focus on physical healthcare only. Consequently, the needs for emotional and mental healthcare applications are gradually increasing.

We present a system developed with the aim of caring for the user's emotions; the proposed system automatically generates and displays the users' portrait drawings by utilizing IoT. In this system, when a user takes photographs throughout the day using a smartphone, the most exciting portrait is selected automatically based on the user's heart rate taken from a smartwatch equipped with the heart rate sensor. The selected photo is then rendered as portrait drawings with various styles by using the nonphotorealistic rendering (NPR) techniques at the server located at the user's home. Finally, when the user returns home, the stylized portrait drawings are displayed in a digital photo frame that is connected to the server. Consequently, the user can recall the most exciting and also the happiest moment of the day through admiring the drawings and heal their emotions accordingly.

The major contributions of this study are as follows. First, we present an IoT-based application for human emotional care as mentioned earlier. This is achieved by employing the NPR techniques, which aim to acquire aesthetic images. The study not only shows that the NPR can be utilized for emotion-aware applications but also that using IoT technique can further expand it. Second, we propose a novel portrait stylization method inspired by Henri de Toulouse-Lautrec's etude, which mostly stylizes the facial region while abstractly representing the others. In this study, our system cares for human emotions using automatically stylized portrait drawings. Selective focusing and expressing the facial region during stylization process, therefore, play a key role in enabling the resulting drawings to convey facial expression

representing the most exciting moment of the day. Finally, by conducting user studies, we show that our system makes users feel better after recalling the most exciting moments of their days by admiring the drawings our system generates. Through the user study, we also confirm that the stylization method proposed in this paper generates results that reasonably resemble Henri de Toulouse-Lautrec's etude.

The remainder of this paper is organized as follows. In Section 2, we provide an overview of related works on IoT applications and artistic stylization methods. Next, we describe our system which automatically and remotely generates portrait drawings using IoT technique in Section 3. We then present our portrait stylization approaches in Section 4. In Section 5, we present experimental results and their evaluation. Finally, in Section 6, we conclude with a summary of our method and the scope for future development.

2. Related Work

Due to the popularization of mobile devices, the range of Internet of Things (IoT) technology has been rapidly broadening. This has made IoT technology nearly ubiquitous in our lives, and it widely supports facets of our well-being. In this regard, many studies focused on various IoT applications. For healthcare, [1, 2] developed applications which monitor heart rate measured by heart rate sensor of smartwatch and provide personalized information for the user's heart care without going to hospital or pharmacy to get doctor's diagnosis. For fitness, applications where activity tracker communicates with fitness facilities like running machine and cycling fitness to obtain precise fitness data, such as exercise distance, speed, and slopeness, and tracks these on smartphone to suggest personalized exercise plan are widely used [3, 4]. However, despite increasing interest in physical heath, studies for emotional care have rarely conducted. In this paper, we aim at developing a system for human emotional care. To do this, we utilize stylization techniques.

Artistic image stylization is a technique that aims at converting an original image into an artistic image with certain style [5]. Early studies on artistic image stylization [6, 7] focused on expressing painterly style by generating brush strokes. In the studies, brush stroke's properties such as direction, size, and texture were mainly managed to stylize original image. Meanwhile, several studies have focused on expressing specific artistic style. Van Gogh's painting styles were imitated by manipulating brush stroke's properties [8, 9], and Seurat's pointillism was simulated as well by employing color juxtaposition theory [10]. These studies were conducted while targeting on mimicking the impressionists' painting style, because the brush strokes they used were quite unique. In this paper, we aim at generating Toulouse-Lautrec's portrait etude who is one of the famous impressionists in 18 centuries.

3. Remote Portrait Generation and Display System for Human Emotional Care

3.1. System Overview. The system proposed in this paper consists of a mobile application part including smartphone application and smartwatch and a server part including server and digital photo frame. In our framework, the user wears a smartwatch and takes portrait photographs throughout the day by using our smartphone application. While taking a photograph, the application also acquires the heart rate from the smartwatch and sends it to the server located at the user's home. The server then stylizes the photographs as portrait drawings with given styles if they are judged as the most exciting moment of the day based on the heart rate. When the user comes back home, the server displays the drawings by using a digital photo frame, and through admiring these drawings and recalling the moments, the user experiences an emotional healing. We present the detail of each part in the next two sections.

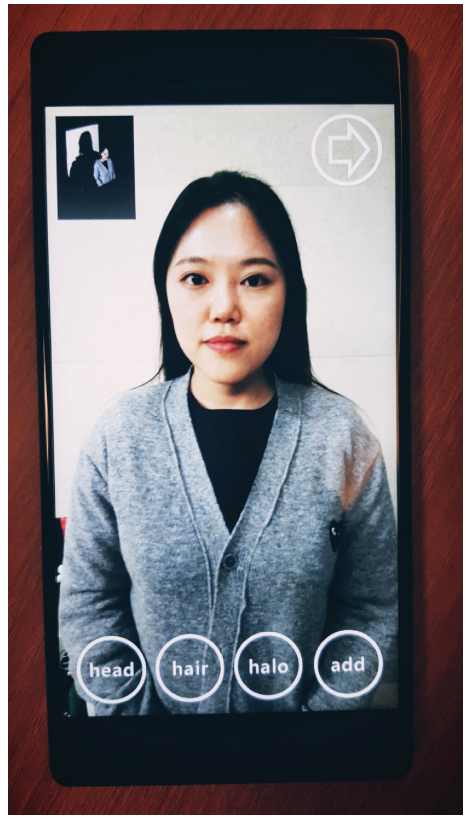


Figure 1: Our smartphone application.

3.2. Mobile Application Part for Acquiring Photograph and Heart Rate. User's photographs are acquired throughout the day by using a smartphone and smartwatch equipped with heart rate sensor, built-in camera, and depth perception sensor. When the user takes a photograph using a smartphone, our smartphone application (Figure 1) simultaneously obtains user's heart rate information from the smartwatch and sends the photograph with heart rate to the server. In this study, we used Samsung Gear S which is a smartwatch equipped with heart rate sensor and integrated it with our smartphone application developed on Android environment.

Our system mainly aims at generating portrait drawings. Consequently, dividing input photograph into various

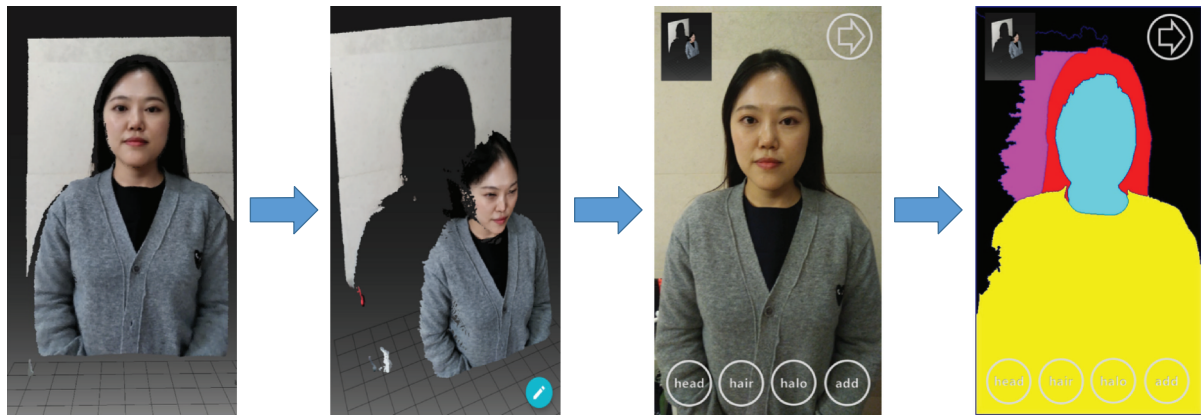


Figure 2: Segmentation process using depth perception.

regions such as facial area and background enables the system to obtain a better quality of the drawings. To do this, we segment the input photograph by using the depth perception of Google's Tango, which is an augmented reality computing platform. Figure 2 shows the process that is performed on Lenovo Phab 2 Pro, which is a smartphone employed in this study and supports the Tango. The Tango reconstructs a 3-dimensional surface from the photograph taken by using depth sensor and motion sensor. Starting from the surface, we search for the point of the greatest depth and find the center of the bounding box, which is detected as the face from the photograph. By using these two points as seeds, we then perform watershed segmentation on depth domain to separate object from the background. Optionally, we also provide semiautomatic segmentation using the seeds mentioned above and user-selected seeds for hair and body. Finally, the segmentation result is incorporated along with the input photograph and is sent to the server to be used in portrait drawing generation process described in the following section.

3.3. Server Part for Generating and Displaying Portrait Drawings. When a photograph is received with its segmentation result and heart rate, the server which is located at the user's home judges whether the photograph is the most exciting moment of the day. In this study, we regard the photograph incorporated with the highest heart rate as the most exciting moment. Therefore, the server renders portrait drawings only if the heart rate of the received photograph is higher than that of previously received photographs. For the style of portrait drawings, we provide oil paintings, pastel drawing, pencil drawings, and portrait etudes, as described in Section 4. Using these four styles, the server generates portrait drawings and preserves them until a new photograph corresponding to a more exciting moment is received. The server monitors the user's location, which is computed on the smartphone using GPS. When the user is within a predefined distance from the server, the server finally displays the portrait drawings on the digital photo frame connected to it.

This paper supposes the system to be used for home use. Therefore, it is recommended to use a low-power device as the server of our system for a practical use. To this end, we used Raspberry Pi 2 B+ and Camel PF1710IPS as the server and digital photo frame, respectively, as shown in Figure 3. When a user admires our resulting drawings, interactions which enable the user to browse and select preferred drawings are provided by an infrared remote controller. To do this, we capture infrared remote signals from the GPIO 22 pin of Raspberry Pi.

4. Portrait Stylization Methods

This study aims at enabling users to admire portrait drawings with various styles to arouse their interests. To achieve this, algorithms for stylizing various styles are required. As mentioned in Section 2, lots of stylization methods have been proposed. Among them, we employ Seo et al.'s painterly rendering algorithm [11] using a stroke texture database. In their study, each individual brush stroke texture, which is captured from a glass sheet where an artist paints on, is stored in their stroke texture database and is utilized for generating brush strokes in the rendering process. At this time, if the database is replaced with another one which consists of brush strokes of different medium, the style of resulting drawing is consequently changed. In this manner, we used three mediums (i.e., oil paint, pastel, and pencil) to provide each style (Figure 4).

In this study, we focus on generating portrait drawings using the most exciting moments taken by users throughout the day. Therefore, conveying facial expression with abstract representation of the others is sometimes more effective for giving a strong impression to the users. In addition to the styles mentioned above, we propose a portrait stylization inspired by Henri de Toulouse-Lautrec's etude, which mostly stylizes the facial region while abstractly representing the others. As shown in Figure 5, Toulouse-Lautrec left many portrait etudes. The main feature of his etude style is represented as simplification except for facial region. In his etudes,

Figure 3: Digital photo frame and server with IR receiver.

(a) (b) (c)

Figure 4: Brush stroke textures: oil paint (a), pastel (b), and pencil (c).

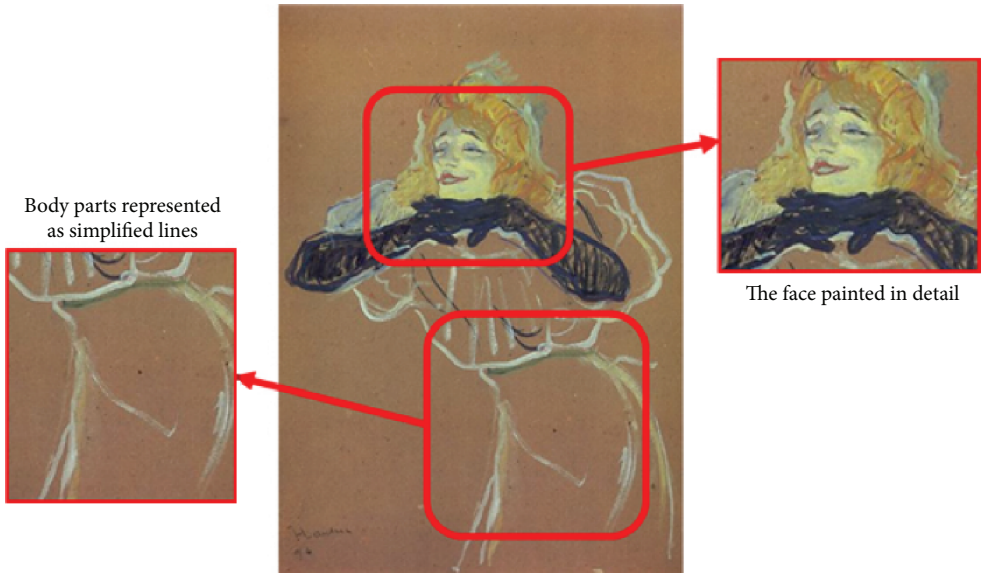


Figure 5: The main feature of Toulouse-Lautrec's portrait etude.

the background is generally eliminated. Only face part is depicted in detail while most of body parts are simplified as a few lines. In this manner, we extract simplified edges for representing body and manipulate them to mimic the lines he used. Then, we paint brush strokes on facial region mostly to make user concentrate on the facial expression. We present the detail of proposed algorithm in the following section.

4.1. Portrait Etude Stylization. Figure 6 shows our portrait etude stylization process. The process is divided into three steps. The first step is line processing. In this step, we extract

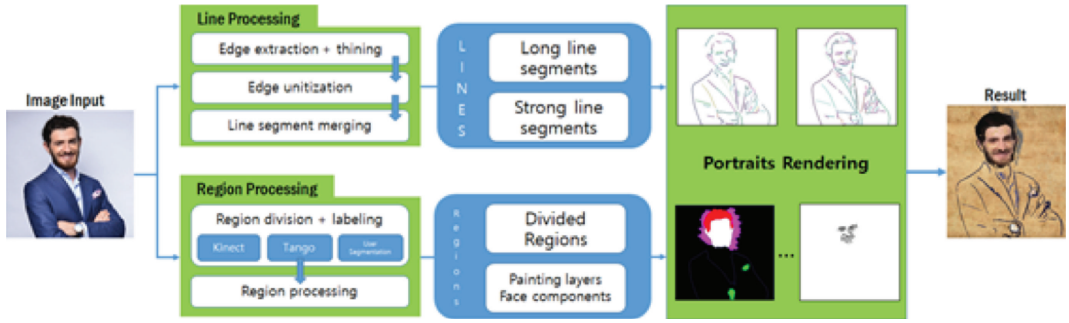


Figure 6: Portrait etude stylization process.

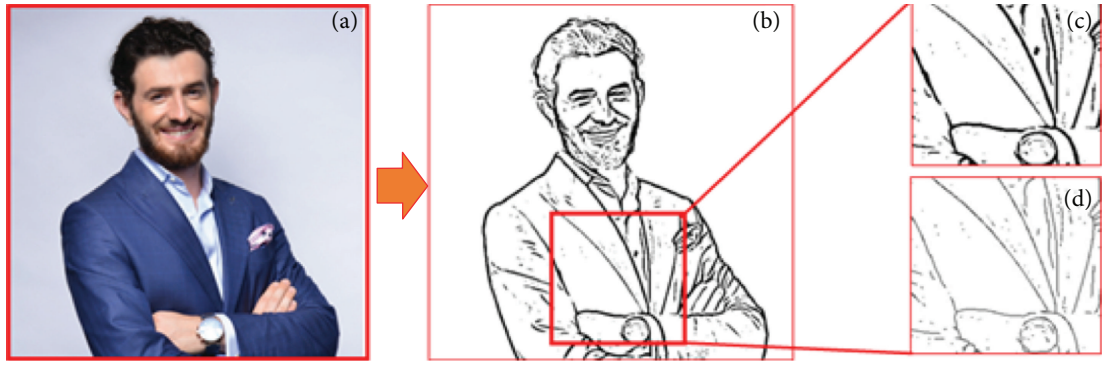


Figure 7: Edges extracted through FDoG (c) and thinned edges (d).

edges from input image and merge them into long simple lines. Next step is region processing; we divide the input image into several regions such as face, hair, body, and background. Final step is rendering processing. In this step, we render line and brush strokes.

4.1.1. Line Processing. In this step, we focus on the expression of simple lines shown throughout Toulouse-Lautrec's portrait etudes, especially for body expression. To generate simplified lines, we first extract edges from input image and vectorize them. In this study, flow-based difference of Gaussian (FDoG) filtering [12] is used to extract edges. As shown in Figure 7, the filtered result is a raster image including thick edges. Therefore, vectorizing edges is necessary to obtain simplified lines. To achieve this, we first use the thinning algorithm to make 1-pixel wide edges (Figure 7(d)). We then use the tracing algorithm to detect and trace each point on edges by using 3×3 kernels as shown in Figure 8. In the figure, the tracing starts from the end point of the edge (a) and proceeds along the edge (b–d). In cases of (a), (b), and (c), there is only one candidate for next edge point which is not traced yet. However, in case of (d), two candidates are given, so choosing one which makes the edge simpler is required. To solve this, we select the point where the angle difference is minimized through the vector calculation (inner product) of the tracked edge points and candidate points.

If more than one pixel is cut off on the edge (Figure 9(a)), it will not be traced to single line (Figure 9(b)) because the kernel covers only adjacent neighbor pixels. To solve this, we merge nearby lines considering the distance and direction between them (Figure 9(c)). If the distance between each end point of two lines and the difference between their directions are below predefined threshold, we merge them into the single line. In this paper, experimentally, thresholds for distance and direction are set to a default value of 5 and 15°, respectively.

Among the separated line segments, we select long and strong line segments only by using their lengths and average gradient values to represent overall structure abstractly. Figure 10 shows the line segments and a histogram which represents their length and numbers. We obtain a threshold value for dividing them into long and short line segments by applying the k-means clustering algorithm with $k = 2$on the histogram. In the figure, long lines (blue) and short lines are divided by using the threshold whose value is set to 32. In the same manner, we divide the line segments into strong and weak lines by using their gradient intensity. In our method, strong lines tend to be detected more in the areas where detailed expression such as face is needed. These two lines are utilized to express each part of input image separately in Section 4.1.3.

4.1.2. Region Processing. To express each region separately according to what the region is, it is needed to divide the input image into regions and label them. As shown in Figure 5, Toulouse-Lautrec's etude can be mainly divided into face, hair, body, and halo in background. As mentioned in Section 3.2, we obtain such regions by using depth perception and user assistance.

In Toulouse-Lautrec's etudes, he briefly depicted facial components such as eyebrows, eyes, nose, and mouth by

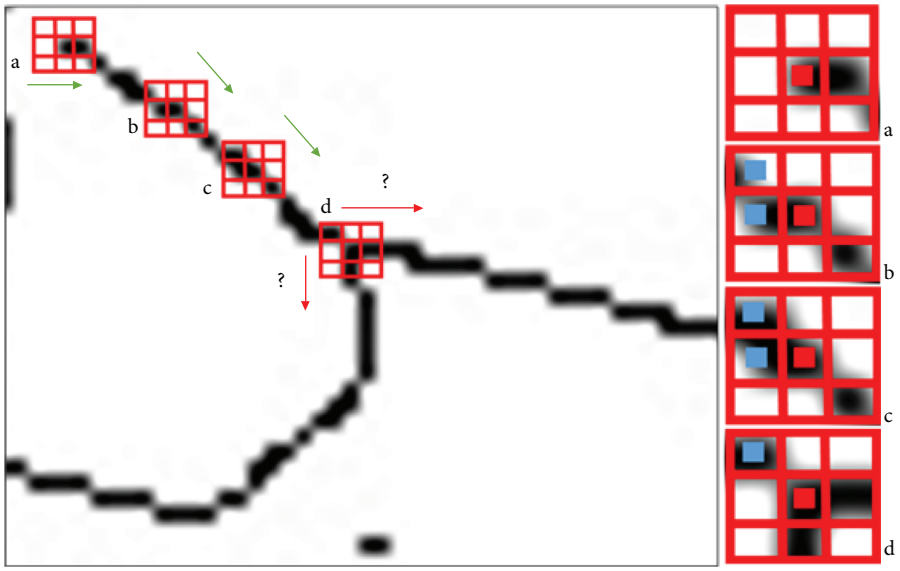
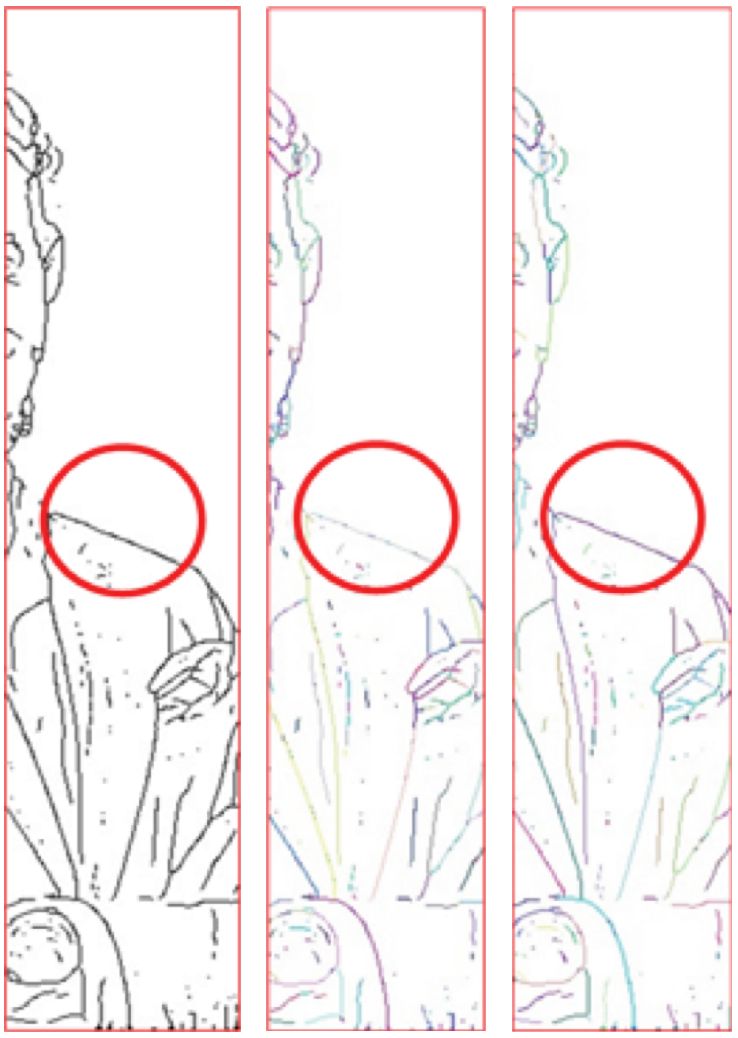


Figure 8: Edge tracing algorithm.



(a) (b) (c)

Figure 9: Line merging: (a) original edge, (b) tracked line, and (c) merged line.

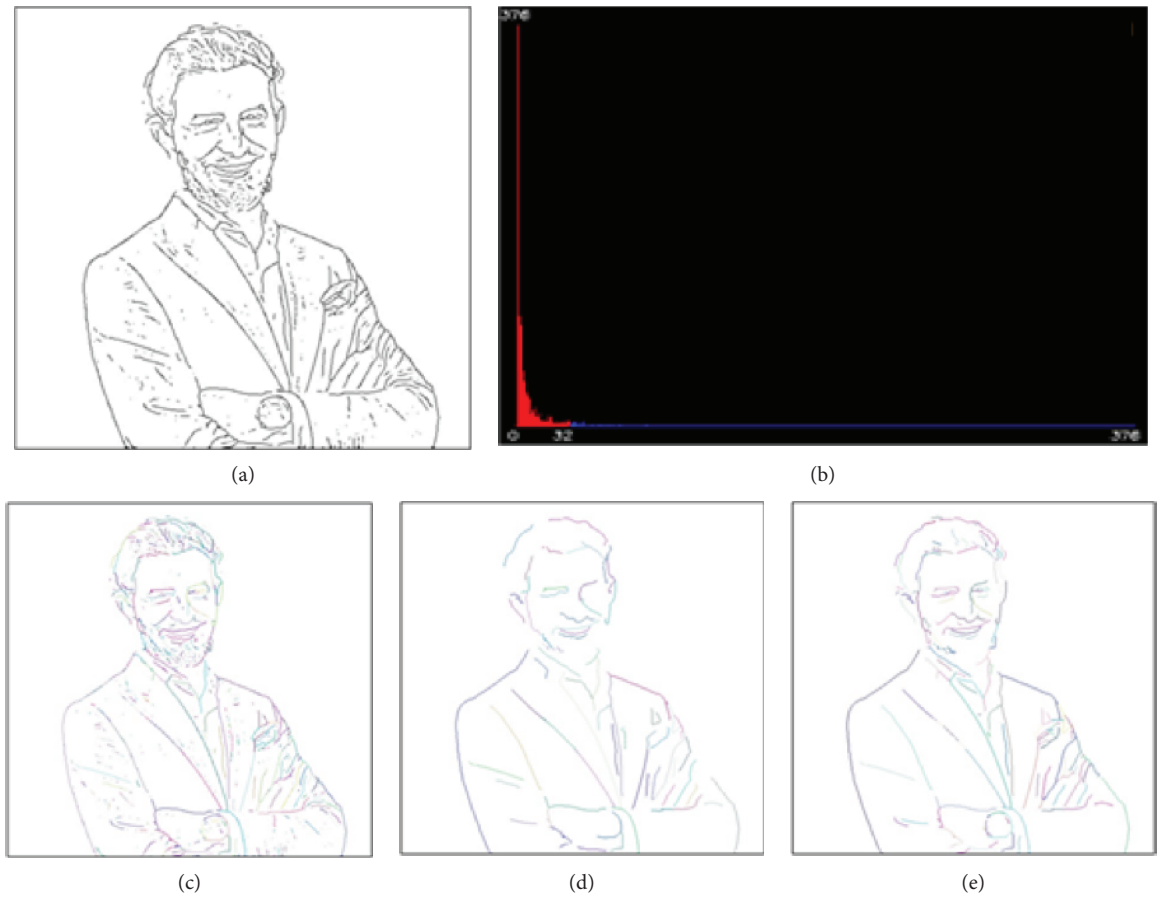


Figure 10: Line selection: (a) original edges, (b) line length histogram, (c) separated line segments, (d) long lines, and (e) strong lines.

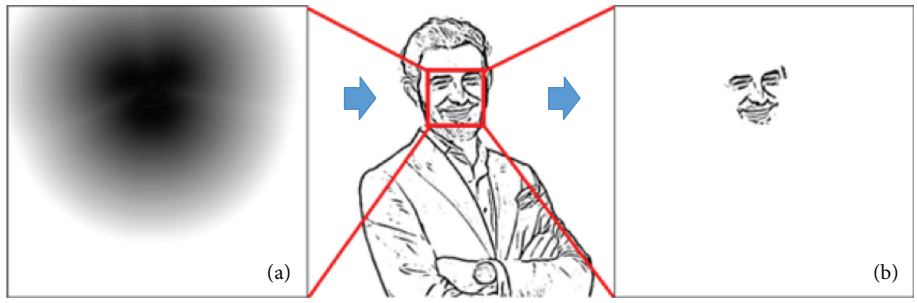


Figure 11: Face edge extraction: (a) facial component distance and (b) face edges.

using simple lines rather than colored brush strokes. To express this, we only draw the lines which correspond to these components. To achieve this, we extract 77 face landmarks from the input image by using active shape model (ASM). We then construct a distance map around facial components which are generated by connecting the landmarks as shown in Figure 11. Finally, we remove out-of-range lines by using the distance map.

For halo region in background, we represent the backlight behind the face. In Toulouse-Lautrec's etudes, the halo is often represented as straight hatching lines with arbitrary direction or following the silhouette of the head. To express this, we calculate modified flow map where the direction is generated by mixing the flow direction obtained by ETF [12] and arbitrary direction as shown in Figure 12.

4.1.3. Rendering Lines and Brush Strokes. Before rendering lines and brush strokes, we create a rectangular background using groundwood paper texture which is shown in Toulouse-Lautrec's etudes. We then render in the order of long lines, halo, face, hair, and facial components. As shown in Figure 13, we render each line through following steps: we first create a triangle strip mesh following the gradient direction; we then color it with the color sampled from input image

(a) (b) (c)

Figure 12: Flow direction modification for halo: (a) original painting gradient, (b) gradient correction map, and (c) modified gradient.

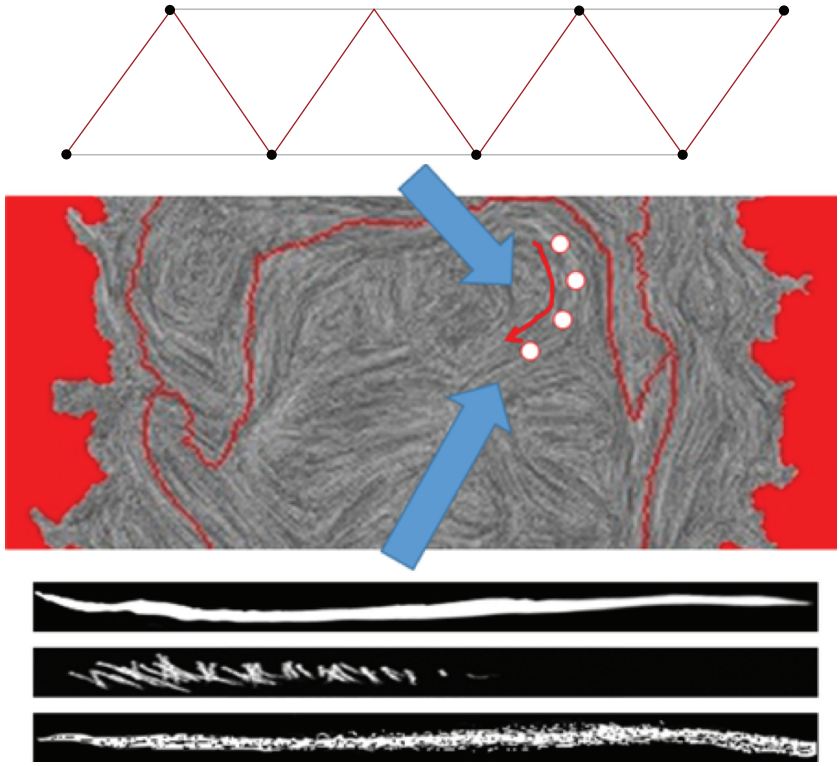


Figure 13: Line rendering using a triangle strip mesh and line texture.

where the mesh is located on; and we finally map predefined line texture captured from Toulouse-Lautrec's etudes.

Our line drawing is performed using the long and strong line segments obtained previously. To render lines with various width, we adjust the width by using the following equation:

$$\text{Thickness} = \frac{\alpha \times \max(w, h)}{\text{mag}}, \quad (1)$$

where w and h are the width and height of the input image, mag is the average gradient strength of the area through which the line segment passes, and α controls the maximum width of lines. Figure 14 shows the lines resulting from adjusting α.

Rendering brush strokes is performed on each region separately. We depict face and hair in detail but body and halo roughly. To do this, we arrange the brush intervals narrow and short on face and hair. On the other hand, for body and halo, the spacing between brushes was widened, and the length was arranged long.

After rendering brush strokes on the entire facial region, we draw facial components defined in Section 4.1.2 to convey facial expression. To achieve this, we first render additional thin brushes on the facial region and then draw the edges corresponding to the facial components. Figure 15 shows the results without and with drawing facial components. As shown in the figure, drawing facial components convey facial expression better.

Figure 16 shows our portrait etude rendering results. As shown in the figure, our method mainly focused on conveying facial expression while representing the others abstractly. In Section 5.2, we evaluate our results by using user study in terms of similarity to Toulouse-Lautrec's etudes.

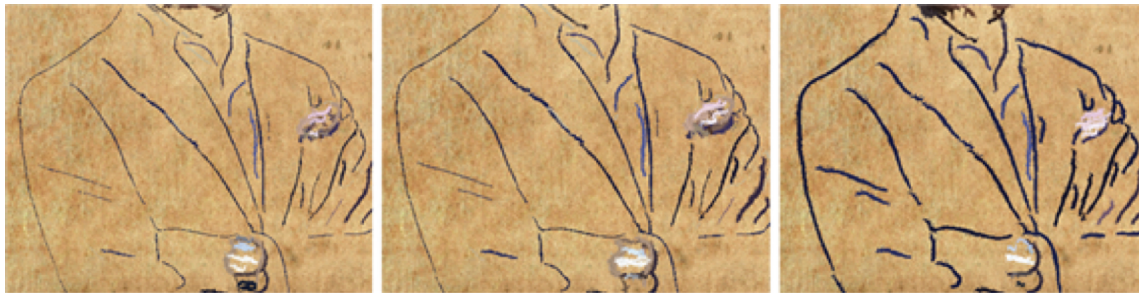


Figure 14: Lines rendered using different user control maximum width.

(a) (b)

Figure 15: Effect of facial components: (a) without facial components and (b) with facial components.

Figure 16: Our portrait etude results.

5. Experimental Results

Figure 17 shows the portrait drawings our system generated. As mentioned above, our system generates four styles of portrait drawings from the most exciting photograph of the day and allows users to select preferred results to display them on the digital photo frame of the system. In the figure, each marked result represents the drawing selected by the user who took the input photograph. In case of that users do not perform the optional semiautomatic segmentation described in Section 3.2, only facial region and background are distinguished; consequently, our portrait etude is not generated. We note that the results which do not contain the portrait etude result correspond to this case. As shown in the figure, the system generated the drawings of the most exciting moments of the day. We can observe that the system rendered facial region in detail compared to the others. It is because our smartphone application divided input photograph into face and the other parts semiautomatically and used them to render each part separately.

In the next two sections, we show that our system makes users feel better after recalling the most exciting moments of their days by admiring the drawings our system generates and also show that our portrait etude stylization

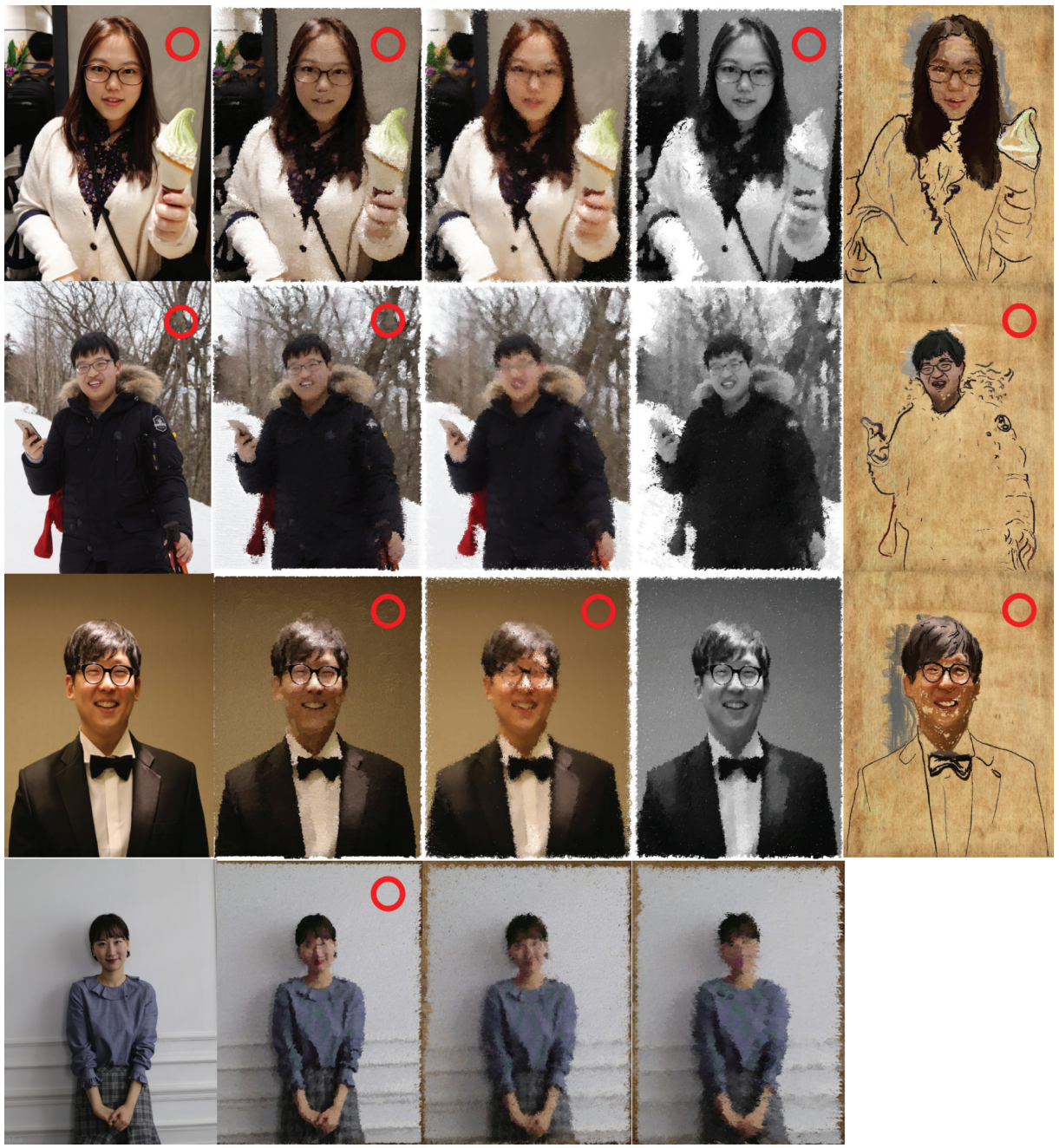


Figure 17: Input photographs and the portrait drawings our system generated. Marked result represents the photograph/drawing selected by the user.

method generates results that reasonably resemble Toulouse-Lautrec's etude.

5.1. Evaluation of the System. We suppose that a user takes photograph throughout the day and uses our system after coming back home where the system is installed. However, we installed the system at an office in the university and conducted a user study. We lent each participant a smartphone where our application was installed and a smartwatch, demonstrated how the devices worked to them, and let them freely take portrait photographs containing themselves throughout the day while wearing the watch. We then made them visit the office, demonstrated how the digital photo frame worked to them, and finally let them admire the original photograph they took and the drawings generated. Participants consisted of 19 undergraduate and graduate students who did not have any expertise in fine art. When they admired the photograph and drawings displayed on digital photo frame, we checked which styles they chose and also checked whether they had performed semiautomatic segmentation after taking the photograph. We then inquired them about the next three questions. (1) Do you think you felt better after admiring the drawings? (2) Do you agree with that the drawings correspond with the most

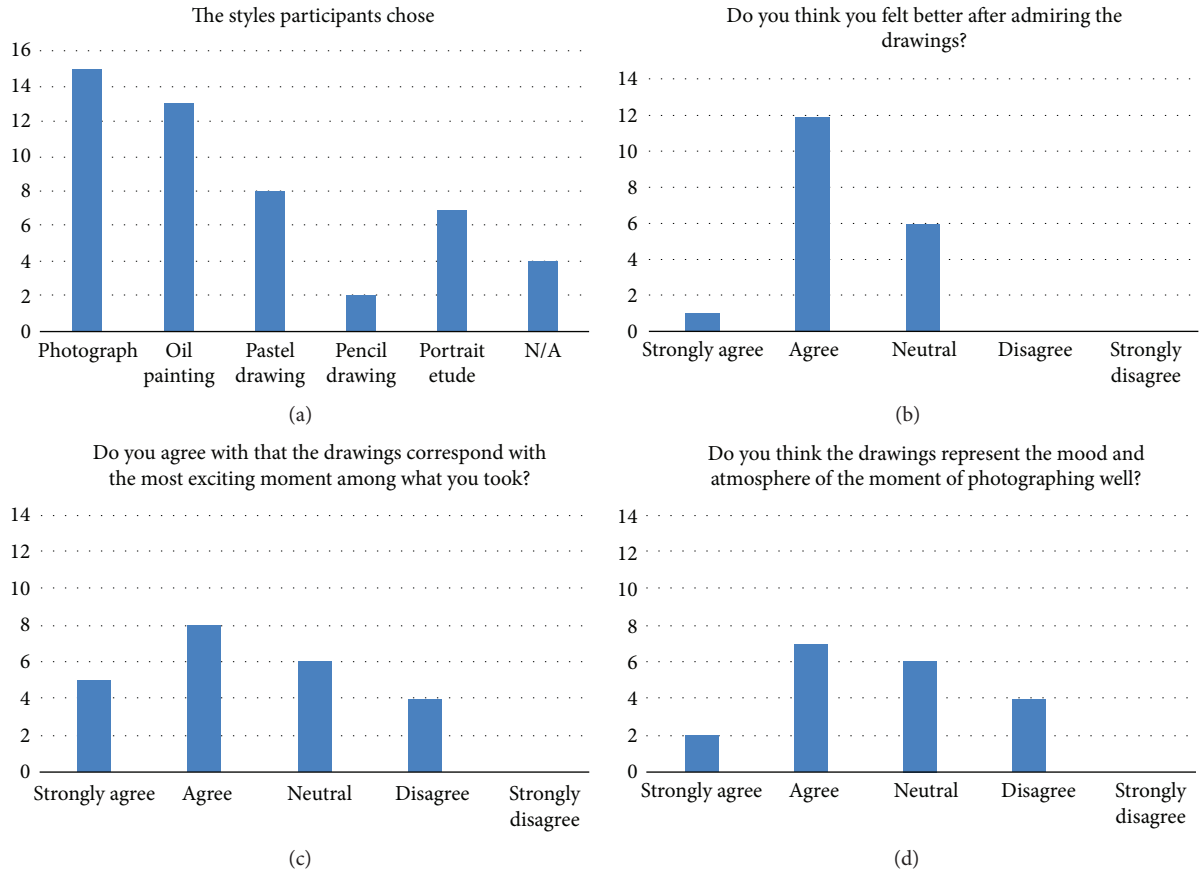


Figure 18: User study results.

exciting moment among what you took? (3) Do you think the drawings represent the mood and atmosphere of the moment of photographing well? For the answers, we use the following scale: 1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, and 5 = strongly agree. In addition, we asked them to give reasons and any feedback.

Figure 18(a) shows the styles participants chose. In the experiments, participants were able to choose multiple styles including original photograph to display them. Otherwise, they were also able to select none of the styles to display all of them. As shown in the figure, among 19 participants, 15 of them chose their preferred styles; the others chose nothing. Regarding the styles chosen by participants, photograph, oil painting, pastel drawing, pencil drawing, and portrait etude were chosen 15, 13, 8, 2, and 7 times, respectively. At this time, we note that portrait etudes were not generated for 7 cases where semiautomatic segmentation was not performed. The result shows that participants mostly preferred original photograph. Among drawings, oil painting was the most preferred style in our system. In this regard, participants responded that original photograph was easy to recognize. They also answered that oil painting was the most familiar style among what we provided. Meanwhile, pencil drawing was chosen only twice. They also mentioned that this was because facial detail was not depicted well in our pencil drawings.

Figure 18(b) shows the answers to question 1. The answers ranged from 3 to 5, and its average value was 3.74. From this result, we confirm that they mostly felt better after admiring the drawings our system generated.

Figure 18(c) shows the answers to question 2. The answers were distributed from 2 to 5, and the average value was 3.37. This was the worst score among our three questions. We assume that this is because our system depends on heart rate to select the most exciting moment. For instance, a participant answered 2 = disagree to Figure 19 of which input photograph was taken after exercise, and we can suppose that heart rate was affected by exercise rather than emotion. In this regard, we present a possible solution for this problem in Section 6.

Lastly, Figure 18(d) shows the answers to question 3. The answers ranged from 3 to 5. Its average value scored 3.95, and this was the best score in our experiments. Especially, the score of participants who chose our portrait etude as their one of preferred styles was 4.42. We analyze that this is because the algorithm effectively conveyed facial expression by focusing on mainly expressing facial region while representing the others abstractly.

Taking all together, we confirm that our system generated portrait drawings which effectively conveyed the mood of the moment of photographing and that users felt better after admiring the drawings, although the drawings did not

Figure 19: The photograph taken after exercise and its drawing results.

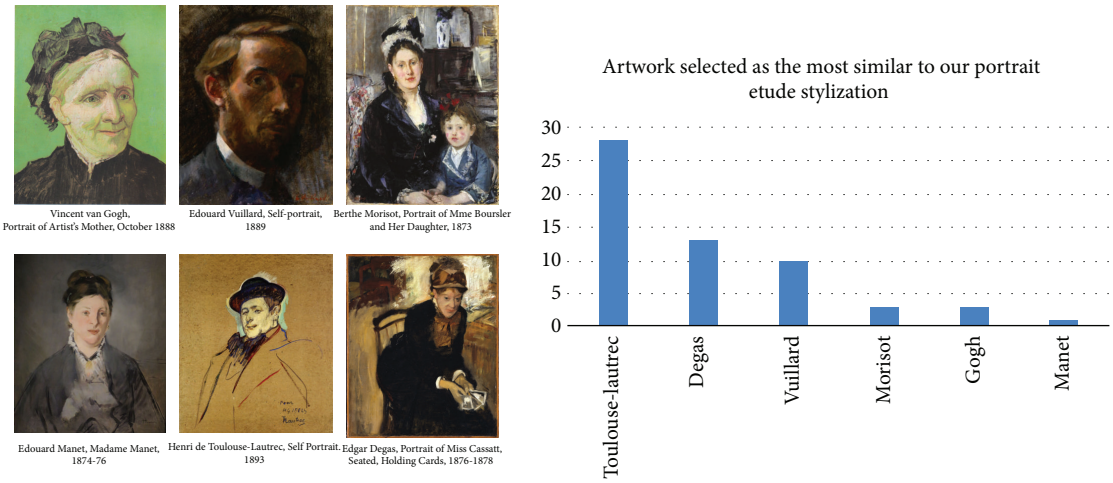


Figure 20: The artworks used for comparison and the number of participants' selections.

well correspond with the most exciting moment of the day. Besides multiple-choice answers, participants' feedbacks were as follows: the styles provided were too limited; portrait etude looked like incomplete; pencil drawing lacked detailed depiction; segmentation process was bothersome; not only the best moment but also the other moments could be displayed; smartphone application drained battery too fast. For these, we also suggest several solutions in Section 6.

5.2. Evaluation of Portrait Etude Rendering. For the style of portrait etude we proposed in Section 4.1, we conducted a user study as well. 31 participants including 19 who participated in the experiment in Section 5.1 consisted of undergraduate and graduate students who did not have expertise in fine art. We showed them 6 portrait artworks (Figure 20) including Toulouse-Lautrec's etude and let them choose the artworks whose styles were similar to our portrait etude drawing results. We allowed multiple-choice questions, consequently 58 selections were made. Figure 20 shows the artworks and the number of participants' selections. As shown in the figure, Toulouse-Lautrec's artwork was selected as the one which was the most similar to our portrait stylization by being chosen by 28 participants. Artworks of Edgar Degas, Eduard Vuillard, Berthe Morisot, Vincent Van Gogh, and Edouard Manet were chosen by 13, 10, 3, 3, and 1 participants, respectively. From this result, we confirm that our portrait etude stylization method generated resulting drawings quite similar to Toulouse-Lautrec's artwork which we targeted. Participants who mainly chose the artworks of Toulouse-Lautrec, Degas, and Vuillard said that line style, halo effect, selective abstraction, and incompleteness were similar to the characteristics found in our resulting drawings. On the other hand, the participants who chose Morisot and Gogh explained that the strokes used for facial area in our results were similar to those of artworks.

We also asked participants how similar Toulouse-Lautrec's artwork and our resulting drawings are. For

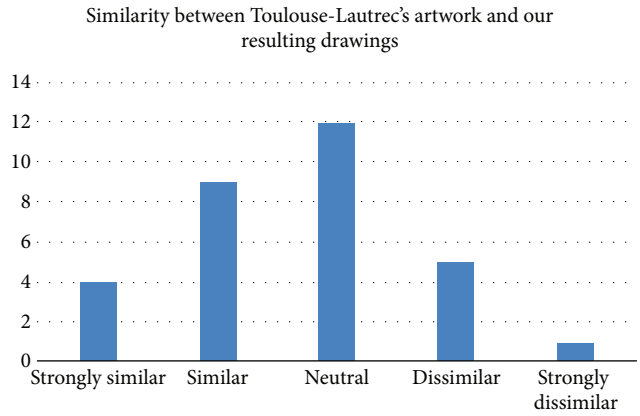


Figure 21: User test result for similarity between Toulouse-Lautrec's artwork and our results.

answers, we used the scale same with Section 5.1. Figure 21 shows the result. The answers were widely distributed from 1 to 5, and the average value was 3.32. Although participants chose Toulouse-Lautrec's artwork as the style most similar to ours, but the score they evaluated was not very high when the result was directly compared with his artwork. In this regard, they mentioned that painting nonfacial parts and using more abstract lines were required to mimic his artwork.

6. Conclusions and Future Work

In this study, we proposed a system that automatically generated portrait drawings for the purpose of human emotional care. In the system, our smartphone application enabled the user to take photographs throughout the day while acquiring heart rates from the smartwatch worn by the user. Meanwhile, the server collected the photographs and heart rates and displayed portrait drawings automatically stylized from the photograph for the most exciting moment of the day. Using the system, the user could recall the exciting and happy moment of the day through admiring the drawings and heal the emotion accordingly. To stylize photographs as portrait drawings, we employed NPR methods, including a portrait etude stylization proposed in this paper. Finally, our system and proposed stylization method were evaluated through user studies.

We collected some ideas for future work from the feedback of user study participants. As mentioned before, participants answered that the drawings our system generated occasionally did not correspond with the most exciting moment of the day. We analyzed that this was caused that our system entirely depended on heart rate to judge it. To solve this problem, we will employ emotion recognition from facial expression. In addition, participants also mentioned that semiautomatic segmentation was bothersome. To provide more better user experiences, we will employ automatic segmentation and labelling methods using semantic recognition. We also believe that this will be beneficial to represent more regions with their own styles to enhance our portrait etude stylization.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

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

Acknowledgments

The authors thank the participants of the user studies of the paper. This work was supported by the National Research Foundation (NRF) of Korea grant funded by the Korea government (MSIT) (no. NRF-2017R1A2B4007481) and Chung-Ang University research grant in 2017.

References

[1] "Hello Heart," http://helloheart.com.

[2] D. Phan, L. Y. Siong, P. N. Pathirana, and A. Seneviratne, "Smartwatch: performance evaluation for long-term heart rate monitoring," in *2015 International Symposium on Bioelectronics and Bioinformatics (ISBB)*, pp. 144–147, Beijing, China, October 2015.

[3] M. Hassanalieragh, A. Page, T. Soyata et al., "Health monitoring and management using Internet-of-Things (IoT) sensing with cloud-based processing: opportunities and challenges. In: services computing (SCC)," in *2015 IEEE International Conference on Services Computing*, pp. 285–292, New York, NY, USA, July 2015.

[4] S. M. R. Islam, D. Kwak, M. H. Kabir, M. Hossain, and K. S. Kwak, "The Internet of Things for health care: a comprehensive survey," *IEEE Access*, vol. 3, pp. 678–708, 2015.

[5] J. E. Kyprianidis, J. Collomosse, T. Wang, and T. Isenberg, "State of the "art": a taxonomy of artistic stylization techniques for images and video," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 5, pp. 866–885, 2013.

[6] P. Haeberli, "Paint by numbers: abstract image representations," in *SIGGRAPH '90 Proceedings of the 17th Annual Conference on Computer Graphics and Interactive Techniques*, vol. 24, pp. 207–214, Dallas, TX, USA, August 1990.

[7] A. Hertzmann, "Painterly rendering with curved brush strokes of multiple sizes," in *SIGGRAPH '98 Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 453–460, New York, NY, USA, 1998.

[8] J. Hays and I. Essa, "Image and video based painterly animation," in *NPAR '04 Proceedings of the 3rd International Symposium on Non-Photorealistic Animation and Rendering*, pp. 113–120, Annecy, France, June 2004.

[9] S. C. Olsen, B. A. Maxwell, and B. Gooch, "Interactive vector fields for painterly rendering," in *GI '05 Proceedings of Graphics Interface 2005*, pp. 241–247, Victoria, British Columbia, May 2005.

[10] S. H. Seo and K. H. Yoon, "Color juxtaposition for pointillism based on an artistic color model and a statistical analysis," *The Visual Computer*, vol. 26, no. 6–8, pp. 421–431, 2010.

[11] S. Seo, S. Ryoo, and P. Jinwann, "Interactive painterly rendering with artistic error correction," *Multimedia Tools and Applications*, vol. 65, no. 2, pp. 221–237, 2013.

[12] H. Kang, S. Lee, and C. K. Chui, "Coherent line drawing," in *NPAR '07 Proceedings of the 5th International Symposium on Non-Photorealistic Animation and Rendering*, pp. 43–50, San Diego, CA, USA, August 2007.

Hindawi
Journal of Sensors
Volume 2018, Article ID 9054758, 10 pages
https://doi.org/10.1155/2018/9054758

Research Article

Biological-Signal-Based User-Interface System for Virtual-Reality Applications for Healthcare

Sang Hun Nam,[1] **Ji Yong Lee**,[2] **and Jung Yoon Kim**[3]

[1]*Department of Newmedia, Seoul Media Institute of Technology, 99 Hwagok-ro 61-gil, Gangseo-gu, Seoul, Republic of Korea*
[2]*HCI and Robotics, University of Science and Technology, 216 Gajeong-ro, Yuseong-gu, Daejeon, Republic of Korea*
[3]*Graduate School of Game, Gachon University, 1342 Seongnam Daero, Sujeong-Gu, Seongnam-Si, Gyeonggi-Do 461-701, Republic of Korea*

Correspondence should be addressed to Jung Yoon Kim; kjyoon79@gmail.com

Received 7 February 2018; Accepted 10 June 2018; Published 29 July 2018

Academic Editor: Banshi D. Gupta

Biosignal interfaces provide important data that reveal the physical status of a user, and they are used in the medical field for patient health status monitoring, medical automation, or rehabilitation services. Biosignals can be used in developing new contents, in conjunction with virtual reality, and are important factors for extracting user emotion or measuring user experience. A biological-signal-based user-interface system composed of sensor devices, a user-interface system, and an application that can extract biological-signal data from multiple biological-signal devices and be used by content developers was designed. A network-based protocol was used for unconstrained use of the device so that the biological signals can be freely received via USB, Bluetooth, WiFi, and an internal system module. A system that can extract biological-signal data from multiple biological-signal data and simultaneously extract and analyze the data from a virtual-reality-specific eye-tracking device was developed so that users who develop healthcare contents based on virtual-reality technology can easily use the biological signals.

1. Introduction

Recently, virtual-reality technology has been effectively used in education, medicine, virtual experiments, and games following the development of hardware and software technologies related to virtual-reality technology [1, 2]. The development of real-time simulations allows perception of virtual reality similar to a real world, and they are utilized in various research studies in combination with user movements [3]. In addition, the theory of games has been successfully used in various other areas other than games because it allows active participation of users by increasing their motivation. Exercise concentration and persistence can be increased by allowing the exerciser to be immersed in the contents by adjusting the difficulty level of the exercise according to the ability of the person [4]. The theory of games applied to health and fitness has been found to influence user behavior changes in real life [5]. Because of the increase in the elderly population and medical expenses, the interest in monitoring health in everyday life and making exercise a part of daily routine has been increasing [6]. It became an exciting area as a social aspect of health and elements of fun and immersion through the combination of virtual- and augmented-reality technologies [7]. Recently, various healthcare contents that measure user movements and exercise methods have been developed because information on the user movements can be obtained following the development of game controllers. Nintendo Wii Fit launched a game that uses body balance using a special controller, and this technology is used in areas such as sense of balance, exercise, corrections, and rehabilitation of the elderly [8, 9]. Microsoft Kinect can be installed in the external environment of a user, and it is used for healthcare monitoring [10] because it can receive data at night using depth sensors and extract user skeletal data. Because the mobile environment in which smartphones are always carried has been developed, various wearable devices have been produced. Users are naturally led to a healthy life by directly obtaining the user's

physical information from wearable devices, updating health information on the smartphone, and measuring the amount of exercise. Recently, interactions with users can be performed through the five senses using virtual-reality technology, and rapid simulation of virtual information has become possible by synchronizing user movements. Because the contents of the applied virtual-reality technology are reinforced with social and physical interaction elements, they are recognized as a new content area and offer the possibility of unlimited development.

Biosignal interfaces provide important data that display the physical status of a user, and they are used not only in the medical field but also various other areas. In the medical field, they are used in monitoring systems for early detection of dangerous situations and diseases by monitoring the patient's health status and in medical automation systems that provide continuous treatment or rehabilitation services. Methods that use biological signals for automatic measurement of stress and objective data collection have achieved practical results [11, 12]. The biological-signal interfaces being used in traditional medical systems are now available for the general public following the development of wearable biological-signal devices. Biological-signal interfaces are used in rehabilitation, fitness, and sports training such as practicing the method of controlling breathing or training body balancing [13, 14]. The application of biological-signal systems has expanded beyond the medical services to various other areas such as education, information security, and human-computer interaction (HCI) as the Internet and mobile devices have become ubiquitous. In the HCI field, research on the experience of users has become important. HCI experts are investigating user content experience, and the industry is measuring the user experience, researching methods of interpreting the collected data, and designing methods of evaluating contents [15, 16]. Measurement of the physiological or physical performance data not only improves our understanding of the physical health but also help us better understand the experience of users by supplementing the results of other methods.

The technology that enables the measurement of biological signals is a fundamental method for creating healthcare contents that regard the biological signal of the user as important. The biological signal is then used as important data that can measure the response of users who use the contents. The present study designed a biological-signal interface system that can be employed in various applications that use biological signals. In particular, considerations were given to special situations employed in healthcare contents in which virtual-reality equipment such as a head-mounted display (HMD), which has been significantly improved in recent years, is used. The rest of the present report is organized as follows. In Section 2, relevant studies that used biological signals are discussed. In Section 3, the architecture and components of the user-interface system are explained. In Section 4, the sensors and data used in the biological-signal interface system are explained. In Section 5, conclusion is made for the proposed system, and new research topics and directions are discussed.

2. Related Works

Biofeedback can be generally classified into direct and indirect. Direct biofeedback is a bodily response such as body or eye movement that can be controlled according to personal intention. Indirect biofeedback is a bodily response such as heart rate (HR) or galvanic skin response (GSR) that cannot be altered by human intervention. Human-behavior analysis technology is widely used for HCI, security surveillance, sports engineering, and intelligent assistance for the elderly. Determining how and what is the manner of movement of people is a fundamental and essential technology to monitor their health. Invasive sensors, which measure human movements, are also used in the field of human activity recognition, and two measurement methods that include a vision-sensor-based method and a method that uses wearable devices are used. Even though vision-based technology reflects the accurate movements of a user, it suffers from a spatial restriction in that external cameras have to be installed. User movements can be measured because inertial sensors can be easily integrated into wearable devices such as smartphones following the recent development of microelectromechanical systems technology. Although the accuracy and location where movements can be measured vary depending on the number and location of sensors, a technology has been developed lately to the level that user motions can be captured [17]. Research on hands and upper limbs or gait analysis for rehabilitation purposes has been conducted [18, 19]. Research on interactive interfaces using gestures for handicapped people has also been carried out. Eye tracking can be helpful for physically handicapped people using the technology for controlling wheelchairs, telepresence, and teleoperation. It has been used in medical and psychological research as a tool to record and study visual behavior [20]. Eye tracking refers to tracking the user's eye movements or the point of gaze. Eye tracking collects important physical information that allows identification of the user intention and is used in wide areas such as psychological research, medical diagnosis, and user and interactive-gaze-control applications [21]. Recent tracking devices are noninvasive and perform tracking based on the image and light reflected from the cornea. In addition to the gaze direction and eye-movement pattern, the measurement data of the pupil size and microsaccades may contribute to the interpretation of the emotional and cognitive status of the user. In contrast to the visual and movement data that change according to personal intention, biological-signal information such as user body temperature and HR can provide not only user health information but also bodily responses that change according to environmental changes [22]. Because noninvasively detecting, collecting, and processing various types of body-related data such as electricity, heat, and optical signals produced by the body have recently become possible, medical expenses can be greatly reduced by making possible the prevention and remote detection of health issues [23]. The extraction of noninvasive biological signals has been interlinked with miniaturized sensor technology and led to the development of wearable devices. It has been integrated to artificial-intelligence technology. The

health-monitoring field is significantly developing. In the past, patient biological signals were monitored and analyzed using equipment installed in the hospital. Many systems that can continuously monitor health have been developed following the development of various types of wearable sensor technology. Because the method of measuring brainwaves through contacts with the head has been sufficiently developed to conveniently collect the brainwave information of a user through wireless electroencephalography headsets, the application of a user interface has been expanding not only to the medical field but also to various other areas, including entertainment and virtual reality [24]. By using biological signals as a user interface, research on controlling a computer cursor by measuring the electromyography (EMG) and electrocardiography (ECG) data of a patient is being conducted [25]. In addition, the sympathetic nerves of the autonomic nervous system increase the HR in response to sudden and strenuous exercises or fear and anger during emergency situations, whereas the parasympathetic nerves decrease the HR during resting state. The autonomic nervous system is closely related to the biological signals of the heartbeat, blood pressure, respiration, emotion, and body temperature.

Virtual-reality technology is leading to the creation of a new healthcare environment through the use of interfaces with medical technology; this is because virtual-reality technology can be used to visualize three-dimensional (3D) data and enable user interaction in various ways by combining auditory, touch, and haptic technology. Healthcare data such as computed tomography (CT) images, magnetic resonance imaging (MRI), cryosection images, and confocal microscopy images are 3D images; thus, these can be analyzed by visualizing them on a monitor [26]. Furthermore, 3D data are easily transmitted to a virtual-reality system and can be visualized and analyzed using 3D-display equipment, such as a head-mounted display (HMD). These 3D data are also utilized in medical education in areas of surgery, dentistry, and nursing, among others, wherein they are used to design interactive environments using various equipment [27]. Virtual-reality technology has also provided a new direction in biomechanics and rehabilitation, as realistic sensory experiences are delivered and scenarios are recreated using virtual-reality-based analysis methods with existing hardware that measures a user's movement [28]; in particular, such technologies are integrated with wearable equipment, IOT technology, artificial intelligence, and cloud services to develop smart healthcare services that can automatically monitor users' health in spaces used by the users such as homes and surrounding environments.

3. Biological User-Interface System Design

Here, we explain the architecture of a biological-signal system and the role and limitation of layers. Figure 1 shows that the architecture has three important layers of sensors, system, and applications. Each layer provides various services according to the application using the system and requirements.

The first layer is composed of various types of sensor devices in which electricity, heat, chemical, and other signals can be detected from the user body. Most of these sensors, for example, ECG and EMG sensors, can directly collect biological signals. In addition, some sensors such as accelerometers collect raw data that can be used to extract information related to health, and by combining the data with the research purpose of the user or other biological signals, the intention from the biological signals can be more accurately predicted. Table 1 lists the data and provides a brief explanation of the sensor supported by the proposed system. Most sensor devices are equipped with microprocessors and functions to convert raw biological-signal data. Devices such as eye-tracking devices process the data while they are connected to a computer. The sensor device layers transmit data to the biological-signal processing system that are wired or wirelessly connected. The second layer is a biological-signal system. To adapt to the research environment where a considerable number of diverse virtual-reality wearable devices are released, the system was designed to add or remove devices that are used to measure different vital signs depending on the purpose of the application. In addition, to preserve the independence of the sensor device layer, the system includes an interface in which two-way data transmission between the network and modules controlling these devices occurs, thus bridging the sensor layer to the biological-signal system.

These systems can be constructed in various forms, such as in smartphone and computer devices, and support network connections such as cellular, IEEE 802.11 wireless protocol, and Bluetooth interfaces [29]. A virtual-reality healthcare application must support wireless communication as well, because it should be able to measure vital signals even when a subject is freely moving around. The biological-signal systems transmit biological signals to the application layer, but it can also process existing information and transmit the processed data. For example, even though the eye-tracking module transmits camera image data, the system provides processed data produced by tracking or measuring the size of the pupil in the image data. The application layer can be classified into a real-time method and a long-term data storage and analysis method according to the method of using the biological signals. The real-time method is used when users' health status is measured and analyzed from the perspective of UI/UX, or when appropriate interactions should be regenerated in real-time in the virtual-reality application based on the user's health status. The long-term data-analysis method allows the creation of various biological-signal-based applications mainly with the application developer collecting information transmitted to a cloud server through the Internet or supported in the form of the Web.

4. Biological System

In the past, biological-signal monitoring was essential only for important situations in which intensive care for the survival and recuperation of a patient was necessary. However, with the advancement in sensor technology, it is now utilized in other fields such as psychology, kinematics, and modern physiology. The present study explains the technology of obtaining and analyzing biological signals that can be used for health management, rehabilitation, and training contents

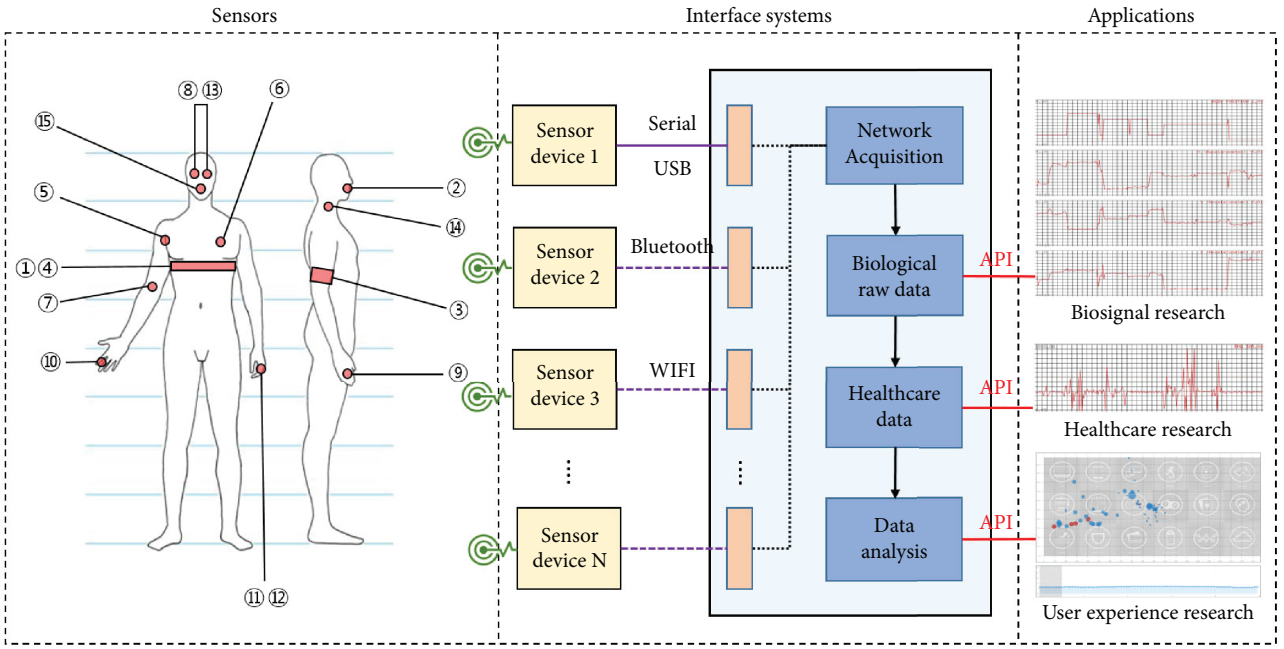


Figure 1: Biological-signal-based user-interface system architecture.

Table 1: Biological signal.

	Sensor	Description of measured components
1	Accelerometer	Three-axis acceleration of the device
2	Airflow	Breathing rate
3	Blood pressure	Systolic, diastolic, and pulse
4	Body position	Five body positions (standing/sitting, supine, prone, left, and right)
5	Body temperature	Body temperature
6	Electrocardiogram	Electrical and muscular function of the heart
7	Electromyogram	Electrical activity of skeletal muscles
8	Eye tracking	Eye tracking and direction of interest
9	Glucometer	Approximate concentration of glucose in the blood
10	GSR	Electrical conductance of the skin
11	HR	Number of heart contractions per minute
12	Oximeter	Arterial oxygen saturation of functional hemoglobin
13	Pupil size	Pupil size of the eye
14	Snore	Snoring
15	Spirometer	Peak expiratory flow

by integrating with the virtual-reality technology. The biological-signal data supported by the proposed system are listed in Table 1, and additional biological signals can be obtained in connection with the system under the provision of network protocols. Here, the medical-use methods through measurement of biological signals supported by the system and the method of using the contents linked to virtual-reality technology are explained.

4.1. Biological Data. The collection and utilization of data used the biological signals. For the biological-signal measurement, "MySignal Hardware Development Platform" was used. Figure 2 shows that 11 biological-signal sensors could

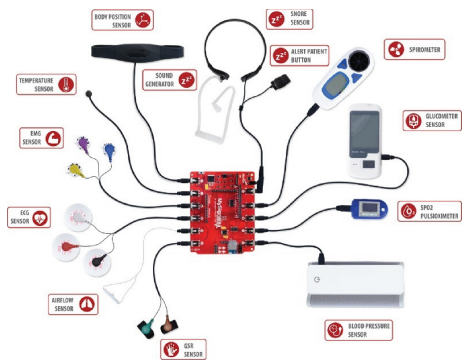


Figure 2: MySignal Hardware Development Platform.

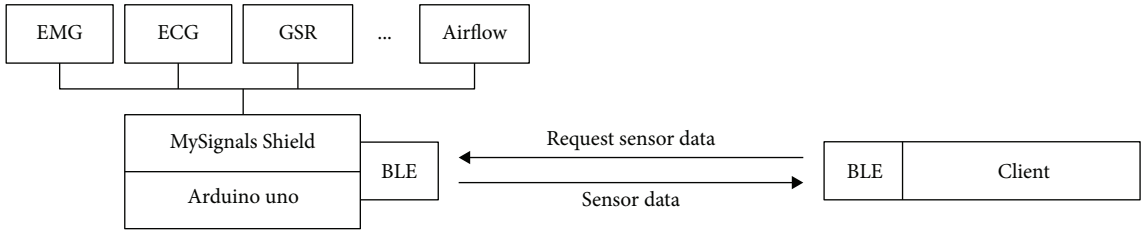


Figure 3: Configuration of a biometric sensor network system.

be used, and the rescue signal from the user could be received. Each sensor was connected using Arduino UNO and MySignals Shield, and the corresponding data were collected. The data collected from each sensor are transmitted to the client-requested sensor data by the client module of PC by serial communication using a Bluetooth low-energy module, as shown in Figure 3. The structure of sensor data is defined differently according to the characteristics of the biological sensor. The client modules are designed and implemented based on the network scheme and specific interface of the biological hardware as shown in Figure 1. However, the client module provides a network-based socket interface so that the same interface can be used on the same system internally or on a remote system connected to the network.

Transmission of an infinite number of sensor data in real time was not possible due to the hardware limitation of Arduino UNO and MySignals Shield. Only the requested data were transmitted by enabling the biological-signal data to activate the sensor required by the client as shown in Figure 4. Thus, transmission of the essential data was ensured and was not interrupted by the transmission of unnecessary data, which prevented wastage in computing resources by unnecessary computation. In addition, the number of bytes of the sensor data being transmitted was minimized. For example, in the case of the GSR sensor, real-number data were obtained using the conductance value. If the data were transmitted as they were, a minimum of four bytes must be transmitted. However, when the data were converted into integer values using a conversion equation, only two bytes were needed to be transmitted. Thus, the transmission efficiency was increased by reducing the amount of data transmitted by converting the real-number data into integers using a conversion rule.

The body-position data enable analysis of the movements caused by specific diseases such as sleep apnea and restless-leg syndrome, and they help determine the sleep patterns through the analysis of movements during sleep. Body-position sensors can help detect the syncope of fall of the elderly and infirm or handicapped people [30]. The data input from the sensors can measure six status ((1) supine, (2) left, (3) right, (4) prone, (5) standing or sitting, and (6) undefined), and by acquiring the acceleration data from a built-in triple-axis accelerometer, additional research can be conducted as shown in Figure 5.

Monitoring the skin and body temperature is medically very important because many diseases accompany changes in the characteristics of body temperature. Body temperature is an important measurement factor for research on such areas as cardiac surgeries and sleep and circadian rhythm and is used for monitoring patient status and predicting risks [31, 32]. As shown in Figure 6, body temperature can be measured using various sensors such as infrared (IR) sensors, thermistors, and thermocouples, and it can be monitored without inconvenience to the user life when a wearable device is used [33].

EMG is a diagnostic technique that measures and records electrical activities produced by skeletal muscles. EMG produces electromyograms by detecting the electrical potentials created by muscle cells when they are electrically or neurologically activated using an electromyograph as shown in Figure 7. By analyzing the electromyogram, biodynamics at the level of medical abnormalities or movements can be analyzed. EMG is used in various clinical and biomedical tests such as diagnosing neuromuscular diseases or as a research tool for exercise science. Because of the development of EMG detection technology, comfortable data collection without attaching electrodes, for example, Myo products, to the body becomes possible. As a contactless interface, numerous applications of the technology are expected [34, 35].

ECG is one of the medical tests most commonly used in modern medicine. Because the measurement of HR in unit of beats using ECG and the exact representation of heart activity in waveforms are possible as shown in Figure 8, accurate monitoring of HR variability is possible [36, 37]. ECG plays a very important role in the diagnosis of cardiac diseases from myocardial ischemia and infarction to syncope and palpitations. Abnormal change in respiration and breathing rate is one of the broad indicators of major physiological instability as shown in Figure 9. Hypoxemia and apnea symptoms can be diagnosed by monitoring the patient condition using the breathing rate.

GSR is a method of measuring the electrical conductivity of the skin, and the result varies depending on the level of moisture as shown in Figure 10. Because sweat glands are controlled by the sympathetic nervous system and a moment of strong emotions changes the electrical resistance of the skin, skin conductivity is used as an indicator of psychological or physiological arousal.

Blood pressure is the pressure of the blood in the arteries when the blood is circulating through the body due to the heartbeat. The blood pressure is recorded in two numbers of systolic pressure and HR. Because high blood pressure can lead to serious problems such as cardiac arrest, cerebral stroke, or renal diseases, it is one of the important physical indicators that must be regularly checked [38]. Pulse

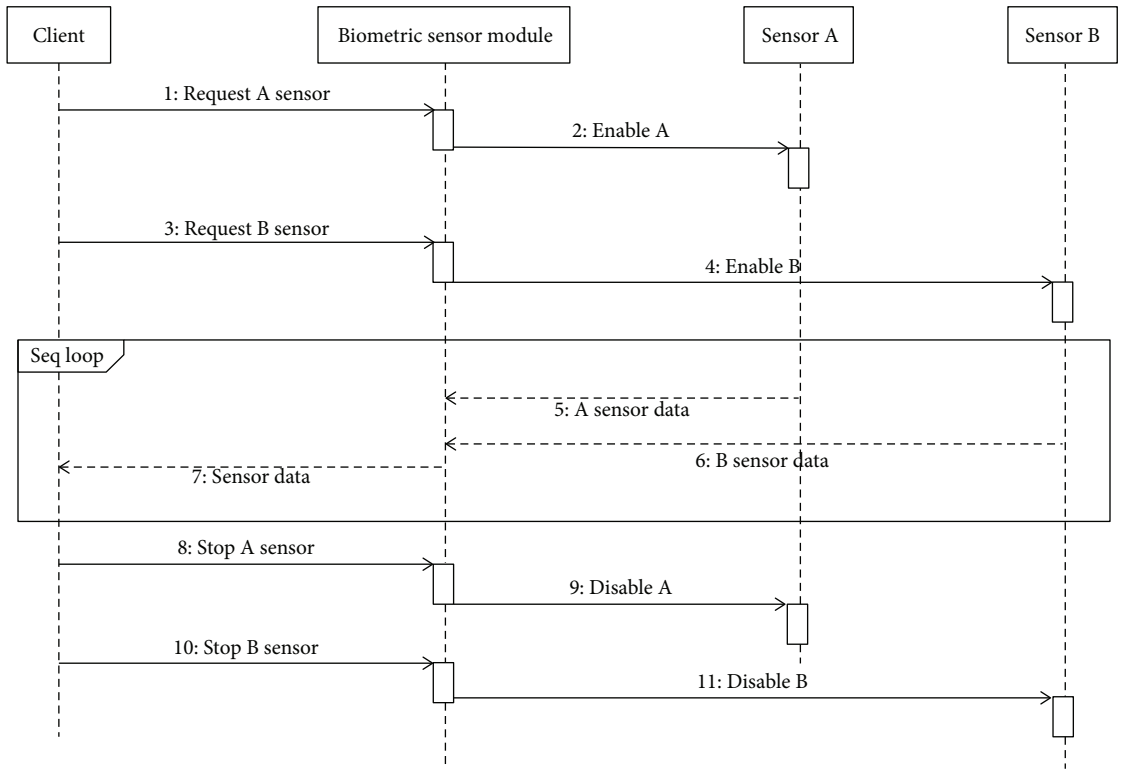


Figure 4: Sequence diagram for biosignal data request.

oximetry is a noninvasive measurement method that shows the arterial oxygen saturation of hemoglobin as shown in Figure 11. It measures the amount of oxygen dissolved in the blood using two waves of absorption coefficient, namely, 660 nm (red light spectrum) and 940 nm (IR light spectrum), based on the detection of hemoglobin and deoxyhemoglobin. Pulse oximeter sensors are required to analyze the treatment, surgery, and recovery of critically ill patients and for assessment of the oxygen uptake and effects of supplemental oxygen on emergency patients. It is less accurate than ECG but effective for long-term monitoring and is widely applied because it can be used in smart watches and wristbands [39].

Glucometer sensors can help manage the blood glucose of diabetic patients by measuring the blood sugar level. Diabetic patients should manage their blood glucose through periodic measurement. Spirometry is the most common pulmonary function test that examines and measures lung function and measures the amount or speed of air that can be inhaled and exhaled. It measures the breathing capacity, which is useful for evaluating conditions such as asthma, fibroid lung, cystic fibrosis, and chronic obstructive pulmonary disease. Snoring is a major symptom of obstructive sleep apnea. To analyze snoring, data must be collected at a high sampling rate, and a large amount of data are processed in most studies because snoring is analyzed using the acoustic characteristics of the sound collected through a microphone. The system designed in the present study employed a hidden Markov model- (HMM-) based analysis method that extracts snoring using a piezo sensor attached to the neck as shown in Figure 12. Short-term Fourier transform and short-term energy are calculated and input to the

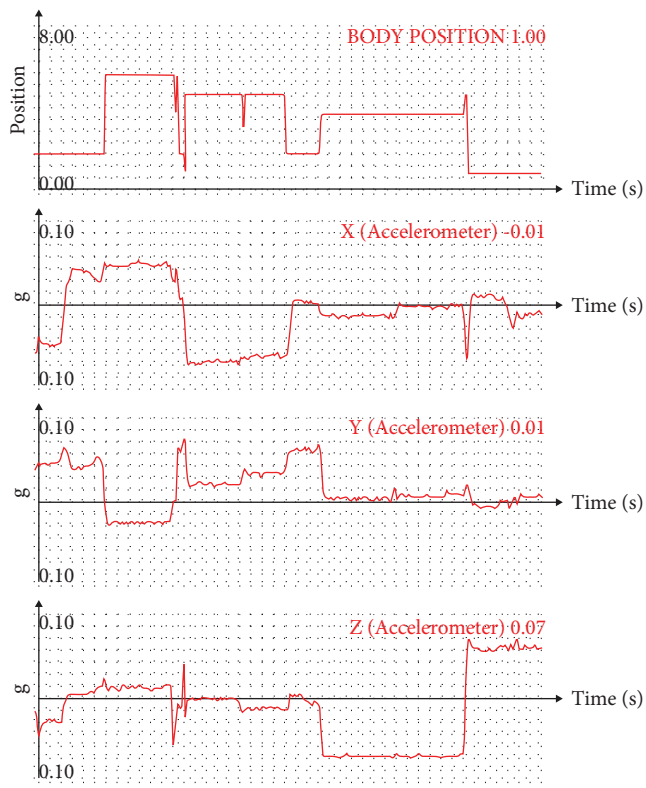


Figure 5: Body-position tracking and 3-axis acceleration signals.

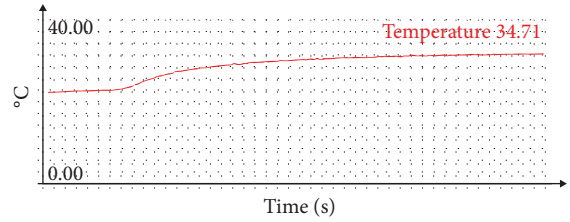


Figure 6: Body temperature signals.

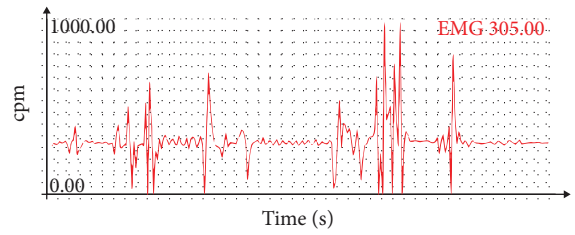


Figure 7: EMG signals.

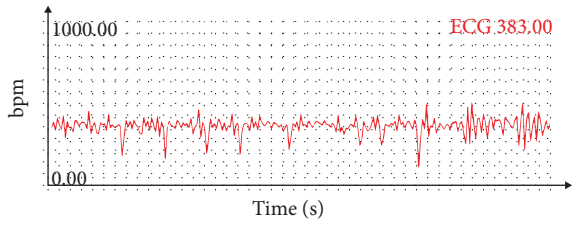


Figure 8: ECG signals.

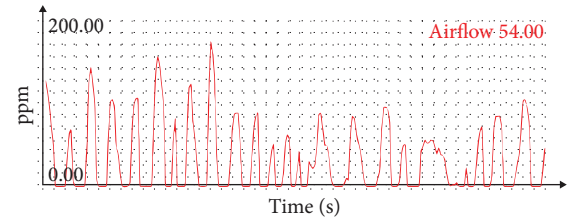


Figure 9: Air flow signal graph.

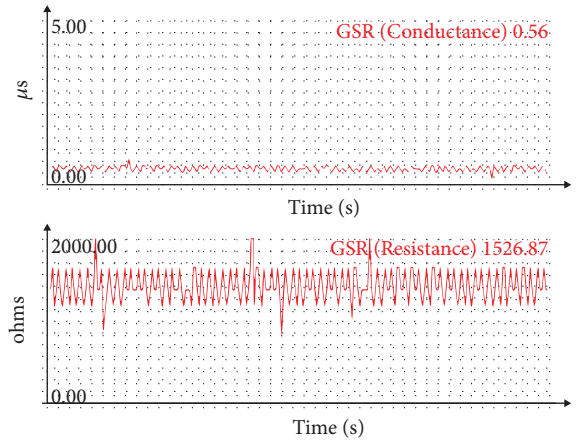


Figure 10: GSR signal graph.

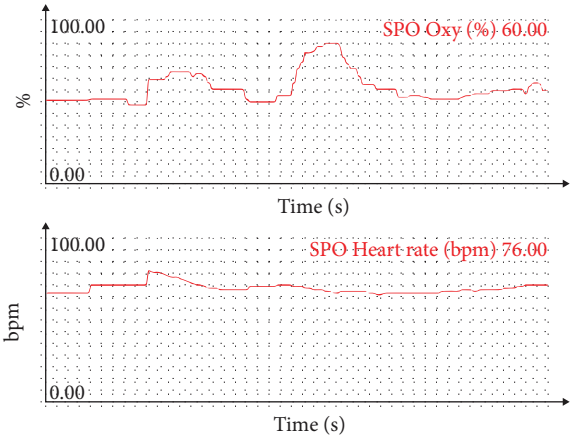


Figure 11: SpO_2 signals.

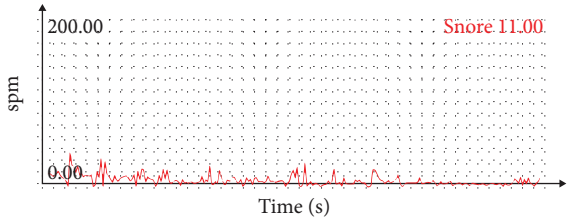


Figure 12: Snoring signals.

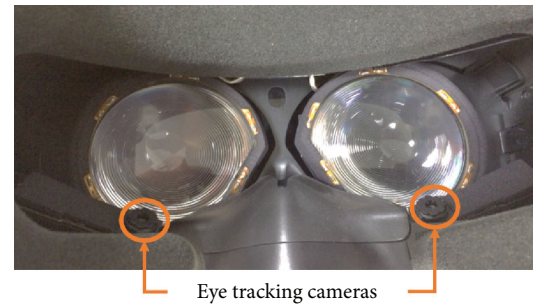


Figure 13: Eye-tracking camera module in HMD.

HMM, and the results can be classified into snoring, noise, and silence. Although it is not a biological-signal measurement, preparation and support for emergency can be ensured by having patients wear an alarm and an emergency button as a pendant or on their wrists.

4.2. Eye-Tracking Data. Assessing human interests and interpreting eye movements as input modes have been proposed as advanced methods of interactions between humans and computers. Previous case studies in Web usability, marketing, medicine, video games, psychology, and neurology suggested that an eye tracker can be used. In the case of virtual-reality contents, eye-tracking modules are installed inside the HMD, as shown in Figure 13, because they use the same display devices as the HMD. Further, because head-mounted eye-tracking devices provide higher accuracy in measuring ocular fixation, saccades, and dilation of the

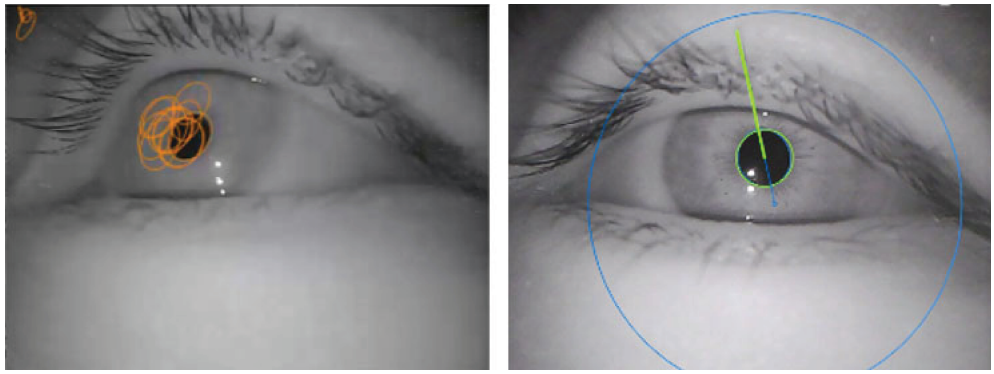


Figure 14: Eye-tracking camera module in HMD.

pupil, they are widely used in psychology and cognitive science research [40].

Figure 14 shows that the pupil of a user is extracted from the IR images transmitted from cameras. The first step in the pupil extraction algorithm finds dark regions in the IR image using binarization. The dark region appears in several places depending on the state of the acquired image inside the HMD. The next step is to extract the contour of candidate regions. Because the extracted contour does not show a perfect pupil shape due to distortion such as image noise, the ellipse fitting is performed using the preliminary information that the pupil is similar to the shape of an ellipse. The small ellipses are removed for eliminating outliers. A heuristic algorithm is applied to determine that the pupil cannot be located at the edge of the image and that the size of the pupil is large and the pupil is located at the center of the image. The pupil data is back projected using the focal length of cameras and extracted ellipse, and the 3D direction vector is calculated through a back-projection process. It is possible to measure the direction of the user's gaze using the 3D pupil data and to monitor the pupil size using the 3D ellipsis of the pupil in real time.

Figure 15 shows the application used to develop the system that tracks eye movements of the user and monitors the photoplethysmogram (PPG) information. A green dot represents the eye tracking of the user on top of the image shown on the HMD. Simultaneously, monitoring of the PPG signal is shown. Figure 16 shows the application that analyzes the biological signals from the user. The top portion shows the changes in the eyeball position transmitted from the user, and the size of the circle represents the changes in the size of the pupil. The bottom portion shows the changes in the user PPG. The user can analyze many biological signals simultaneously extracted under a specific condition through the timeline function.

Figure 15: Eye tracking and PPG capturing application.

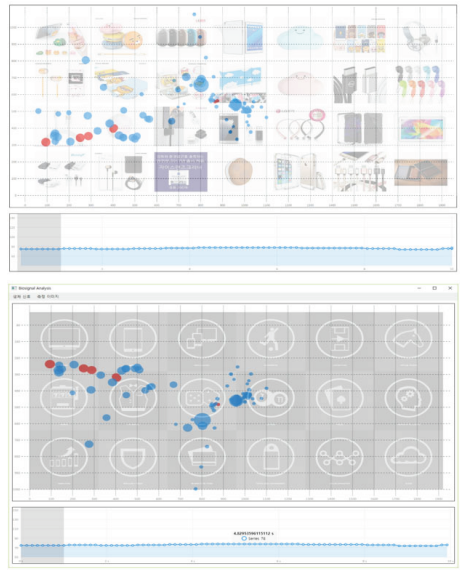


Figure 16: Eye tracking and PPG analysis application.

5. Conclusion

Biological-signal-based systems are widely distributed, along with the distribution of wearable devices, and the range of their applications is also wide. Recent biological-signal modules only support wearable devices supplied by each company, and the trend is not to provide researchers with real-time biological-signal data that could ensure and protect data. Investigators conducting research using biological-signal data can control biological-signal hardware by themselves, but controlling biological-signal hardware is not easy for people creating contents. A biological-signal-based user-interface framework was designed, developed, and tested for those who want to create contents using biological signals. This system supports the researchers that have investigated the relationships among many biological signals using various sensor devices that complement each other and require

analysis of biological signals because it is designed to support many biological hardware devices using a network-based internal interface. Especially, this system can be extended to include the creation of virtual-reality contents that are influencing the medical field, various forms of entertainments, and the education field.

Immersive virtual-reality contents maximize user involvement and users can participate while they are moving around, and new forms of contents can be developed using the biological-signal-based user-interface system. Biological-signal feedback can be obtained from the user in real time while the user is participating in contents, and the learning effect of the user can be increased by changing the contents after analyzing elements of the contents and evaluating biological signals. This study provides a new opportunity of understanding the interaction between humans and computers by designing a system that provides biological signals and attempts to apply user experience by analyzing the relationship between the user's biological signals and emotions in a future study. The traditional method of user evaluation has been carried out using interviews or questionnaires. Since the contents of virtual reality are carried out while users are moving around with the HMD worn, the sense of visual and behavioral immersion increases and a greater amount of concentration and physical strength may be consumed. Since there are large psychological and physical differences between before and after carrying out the contents of virtual reality, it can affect the evaluation process. An analysis on user evaluation using the biological signals of the user can be performed while the contents are being carried out.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

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

Acknowledgments

This research is supported by the Ministry of Culture, Sports, and Tourism (MCST) and the Korea Creative Content Agency (KOCCA) in the Culture Technology (CT) Research and Development Program 2017 (R2017050041_00000001).

References

[1] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, "From game design elements to gamefulness: defining "gamification"," in *Proceedings of the 15th International Academic MindTrek Conference on Envisioning Future Media Environments—MindTrek '11*, pp. 9–15, Tampere, Finland, 2011.

[2] D. Valtchanov, *Physiological and Affective Responses to Immersion in Virtual Reality: Effects of Nature and Urban Settings (M.S. thesis)*, University of Waterloo, 2010.

[3] G. Riva, "From toys to brain: virtual reality applications in neuroscience," *Virtual Reality*, vol. 3, no. 4, pp. 259–266, 1998.

[4] P. T. Chua, R. Crivella, B. Daly et al., "Training for physical tasks in virtual environments: Tai Chi," in *IEEE Virtual Reality, 2003. Proceedings*, pp. 87–94, Los Angeles, CA, USA, 2003.

[5] C. Lister, J. H. West, B. Cannon, T. Sax, and D. Brodegard, "Just a fad? Gamification in health and fitness apps," *JMIR Serious Games*, vol. 2, no. 2, article e9, 2014.

[6] Y. Hao and R. Foster, "Wireless body sensor networks for health-monitoring applications," *Physiological Measurement*, vol. 29, no. 11, pp. R27–R56, 2008.

[7] T. Althoff, R. W. White, and E. Horvitz, "Influence of Pokémon Go on physical activity: study and implications," *Journal of Medical Internet Research*, vol. 18, no. 12, article e315, 2016.

[8] M. Agmon, C. K. Perry, E. Phelan, G. Demiris, and H. Q. Nguyen, "A pilot study of Wii Fit exergames to improve balance in older adults," *Journal of Geriatric Physical Therapy*, vol. 34, no. 4, pp. 161–167, 2011.

[9] G. Saposnik, R. Teasell, M. Mamdani et al., "Effectiveness of virtual reality using Wii gaming technology in stroke rehabilitation: a pilot randomized clinical trial and proof of principle," *Stroke*, vol. 41, no. 7, pp. 1477–1484, 2010.

[10] E. E. Stone and M. Skubic, "Fall detection in homes of older adults using the Microsoft Kinect," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 1, pp. 290–301, 2015.

[11] J. Bakker, M. Pechenizkiy, and N. Sidorova, "What's your current stress level? Detection of stress patterns from GSR sensor data," in *2011 IEEE 11th International Conference on Data Mining Workshops*, pp. 573–580, Vancouver, BC, Canada, December 2011.

[12] D. Cho, J. Ham, J. Oh et al., "Detection of stress levels from biosignals measured in virtual reality environments using a kernel-based extreme learning machine," *Sensors*, vol. 17, no. 10, p. 2435, 2017.

[13] T. Wollmann, F. Abtahi, A. Eghdam et al., "User-centred design and usability evaluation of a heart rate variability biofeedback game," *IEEE Access*, vol. 4, pp. 5531–5539, 2016.

[14] K. M. V. McConville and S. Virk, "Evaluation of an electronic video game for improvement of balance," *Virtual Reality*, vol. 16, no. 4, pp. 315–323, 2012.

[15] D. Clarke and P. R. Duimering, "How computer gamers experience the game situation: a behavioral study," *Computers in Entertainment*, vol. 4, no. 3, p. 6, 2006.

[16] T. Marsh, K. Yang, C. Shahabi et al., "Automating the detection of breaks in continuous user experience with computer games," in *CHI '05 Extended Abstracts on Human Factors in Computing Systems—CHI '05*, pp. 1629–1632, Portland, OR, USA, April 2005.

[17] B. Klaassen, B. J. F. van Beijnum, J. P. Held et al., "Usability evaluations of a wearable inertial sensing system and quality of movement metrics for stroke survivors by care professionals," *Frontiers in Bioengineering and Biotechnology*, vol. 5, 2017.

[18] A. Muro-De-La-Herran, B. Garcia-Zapirain, and A. Mendez-Zorrilla, "Gait analysis methods: an overview of wearable and non-wearable systems, highlighting clinical applications," *Sensors*, vol. 14, no. 2, pp. 3362–3394, 2014.

[19] L. Bai, M. G. Pepper, Y. Yan, S. K. Spurgeon, M. Sakel, and M. Phillips, "Quantitative assessment of upper limb motion in neurorehabilitation utilizing inertial sensors," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, no. 2, pp. 232–243, 2015.

[20] S. I. Ktena, W. Abbott, and A. A. Faisal, "A virtual reality platform for safe evaluation and training of natural gaze-based wheelchair driving," in *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, pp. 236–239, Montpellier, France, April 2015.

[21] P. Majaranta and A. Bulling, "Eye tracking and eye-based human-computer interaction," in *Advances in Physiological Computing*, Springer, London, UK, 2014.

[22] G. G. Berntson, J. T. Bigger Jr., D. L. Eckberg et al., "Heart rate variability: origins, methods, and interpretive caveats," *Psychophysiology*, vol. 34, no. 6, pp. 623–648, 1997.

[23] S. C. Mukhopadhyay, "Wearable sensors for human activity monitoring: a review," *IEEE Sensors Journal*, vol. 15, no. 3, pp. 1321–1330, 2015.

[24] Y. Liu, O. Sourina, and M. K. Nguyen, "Real-time EEG-based human emotion recognition and visualization," in *2010 International Conference on Cyberworlds*, pp. 262–269, Singapore, Singapore, October 2010.

[25] A. B. Barreto, S. D. Scargle, and M. Adjouadi, "A practical EMG-based human-computer interface for users with motor disabilities," *Journal of Rehabilitation Research and Development*, vol. 37, no. 1, pp. 53–63, 2000.

[26] B. Preim and D. Bartz, *Visualization in Medicine: Theory, Algorithms, and Applications*, Morgan Kaufmann Publishers Inc., 2007.

[27] K. M. Hanson, *The Utilization of Mixed-Reality Technologies to Teach Techniques for Administering Local Anesthesia (M.S. thesis)*, Utah State University, 2011.

[28] A. Garcia, N. Andre, D. Bell Boucher, A. Roberts-South, M. Jog, and M. Katchabaw, "Immersive augmented reality for Parkinson disease rehabilitation," in *Virtual, Augmented Reality and Serious Games for Healthcare 1*, pp. 445–469, Springer, Berlin, Heidelberg, 2014.

[29] H. Alemdar and C. Ersoy, "Wireless sensor networks for healthcare: a survey," *Computer Networks*, vol. 54, no. 15, pp. 2688–2710, 2010.

[30] N. Noury, T. Herve, V. Rialle et al., "Monitoring behavior in home using a smart fall sensor and position sensors," in *1st Annual International IEEE-EMBS Special Topic Conference on Microtechnologies in Medicine and Biology. Proceedings (Cat. No. 00EX451)*, pp. 607–610, Lyon, France, 2000.

[31] W. Chen, S. Dols, S. B. Oetomo, and L. Feijs, "Monitoring body temperature of newborn infants at neonatal intensive care units using wearable sensors," in *Proceedings of the Fifth International Conference on Body Area Networks—BodyNets '10*, pp. 188–194, Corfu, Greece, September 2010.

[32] W. Vanmarkenlichtenbelt, H. Daanen, L. Wouters et al., "Evaluation of wireless determination of skin temperature using iButtons," *Physiology & Behavior*, vol. 88, no. 4-5, pp. 489–497, 2006.

[33] A. Pentland, "Healthwear: medical technology becomes wearable," *Computer*, vol. 37, no. 5, pp. 42–49, 2004.

[34] E. Kutafina, D. Laukamp, and S. M. Jonas, *Wearable Sensors in Medical Education: Supporting Hand Hygiene Training with a Forearm EMG*, pHealth, 2015.

[35] J. Nishida, K. Takahashi, and K. Suzuki, "A wearable stimulation device for sharing and augmenting kinesthetic feedback," in *Proceedings of the 6th Augmented Human International Conference—AH '15*, pp. 211-212, Singapore, Singapore, March 2015.

[36] D. Phan, L. Y. Siong, P. N. Pathirana, and A. Seneviratne, "Smartwatch: performance evaluation for long-term heart rate monitoring," in *2015 International Symposium on Bioelectronics and Bioinformatics (ISBB)*, pp. 144–147, Beijing, China, October 2015.

[37] M. Weippert, M. Kumar, S. Kreuzfeld, D. Arndt, A. Rieger, and R. Stoll, "Comparison of three mobile devices for measuring R-R intervals and heart rate variability: Polar S810i, Suunto t6 and an ambulatory ECG system," *European Journal of Applied Physiology*, vol. 109, no. 4, pp. 779–786, 2010.

[38] A. Pantelopoulos and N. G. Bourbakis, "A survey on wearable sensor-based systems for health monitoring and prognosis," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 40, no. 1, pp. 1–12, 2010.

[39] J. Parak, A. Tarniceriu, P. Renevey, M. Bertschi, R. Delgado-Gonzalo, and I. Korhonen, "Evaluation of the beat-to-beat detection accuracy of PulseOn wearable optical heart rate monitor," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 8099–8102, Milan, Italy, August 2015.

[40] L. A. Granka, T. Joachims, and G. Gay, "Eye-tracking analysis of user behavior in WWW search," in *Proceedings of the 27th Annual International Conference on Research and Development in Information Retrieval—SIGIR '04*, pp. 478-479, Sheffield, UK, July 2004.

Hindawi
Journal of Sensors
Volume 2018, Article ID 5034684, 10 pages
https://doi.org/10.1155/2018/5034684

Research Article

Landmark-Guided Local Deep Neural Networks for Age and Gender Classification

Yungang Zhang[1] **and Tianwei Xu**[2]

[1]*Department of Computer Science, Yunnan Normal University, Kunming, Yunnan 650500, China*
[2]*Graduate School, Yunnan Normal University, Kunming, Yunnan 650500, China*

Correspondence should be addressed to Yungang Zhang; yungang.zhang01@gmail.com

Received 22 February 2018; Accepted 30 April 2018; Published 9 July 2018

Academic Editor: Mucheol Kim

Many types of deep neural networks have been proposed to address the problem of human biometric identification, especially in the areas of face detection and recognition. Local deep neural networks have been recently used in face-based age and gender classification, despite their improvement in performance, their costs on model training is rather expensive. In this paper, we propose to construct a local deep neural network for age and gender classification. In our proposed model, local image patches are selected based on the detected facial landmarks; the selected patches are then used for the network training. A holistical edge map for an entire image is also used for training a "global" network. The age and gender classification results are obtained by combining both the outputs from both the "global" and the local networks. Our proposed model is tested on two face image benchmark datasets; competitive performance is obtained compared to the state-of-the-art methods.

1. Introduction

Age estimation and gender distinction from face images play important roles in many computer vision-based applications, such as visual surveillance, security control, and human-computer interaction. Over the last decades, many methods have been proposed to tackle the age and gender classification task.

In early works, pixel intensity values are used directly as input to train a classifier such as neural network [1, 2] or support vector machine (SVM) [3]. However, when the resolutions of images increase, directly using intensity values dramatically increases the scales of image features as well. Therefore, some feature reduction techniques such as principal component analysis (PCA) are applied to reduce the dimensions of image features [4]. Some image descriptors which are more powerful for image representation have also been used in the area of age estimation and gender recognition tasks, such as local binary patterns (LBP) [5], shift-invariant feature transform (SIFT) [6], Gabor filters [7], histogram of oriented gradient (HOG) [8], and biologically inspired features (BIF) [9]. Although the tasks of age estimation and gender classification have been widely investigated over the last decades, the results obtained are still far away from real applications [10, 11].

In recent years, deep learning, especially convolutional neural networks (CNN) [12–15], have become an important tool in computer vision applications. In many vision-based areas, such as image classification, object detection, pose estimation, visual tracking, CNN have achieved superior results. [16]. More recently, CNN have been employed in face image-based age and gender classification tasks [17–19]. However, as face images vary in a wide range under the unconstrained conditions (namely, in the wild), the performances of CNN still need to be improved, especially in age estimation tasks. Moreover, the time cost on training CNN models is quite expensive in most proposed solutions.

In order to reduce the cost on CNN model training, a local deep neural network (LDNN) was proposed [20] for gender recognition; the LDNN model can achieve state-of-the-art performance while the training cost is considerably reduced. More recently, a modified version of LDNN is

proposed by Liao et al. [21]; this modified LDNN shares the same network architecture with the one used in [20]. In [21], the number of image patches used for network training is further reduced; 9 fixed image patches are selected for network learning. The modified LDNN model can be used for both age and gender recognition. However, in this model, the local image patches used for training are fixed; this may not work well on the unconstrained images without carefully preprocessing. In [21], the authors find the eye areas and the mouth area are crucial parts for age estimation, while only the eye areas are important for gender classification.

The success of LDNN in age and gender classification and the relative discoveries from the former LDNN works inspire us to propose a LDNN model for age and gender estimation. In our proposed model, the local image patch selection is based on the detected facial landmarks, that is, the image patches used for network learning are dynamically generated. Therefore, the number of image patches can be greatly reduced while all the important information in a face image can be kept.

In [20], the Sobel edge detector is used for local feature extraction. However, in [21], it is illustrated that using other feature detectors can obtain different performances. In our proposed model, the holistically-nested edge detection (HED) [22] is used for global feature extraction. The age and gender classification results are obtained by combining both the outputs from the "global" and the local networks.

The remainder of the paper is listed as follows. In Section 2, a brief review of related work on age and gender classification using CNN is given. Section 3 introduces the proposed local deep neural network for age and gender estimation. Section 4 presents the experiment settings and the experimental results and analysis. Conclusions and future work are included in Section 5.

2. Related Work

The successful applications of CNN on many computer vision tasks have revealed that CNN is a powerful tool in image learning. If enough training data are given, CNN is able to learn a compact and discriminative image feature representation. Therefore, many researchers propose to use CNN in age and gender classification from face images. In this section, the related work on age and gender classification using CNN is briefly reviewed. The previous research on local deep neural networks for age and gender estimation is also introduced.

2.1. CNN for Age and Gender Estimation. An early CNN model used for age and gender estimation can be seen in [23], in which a multiscale convolution neural network model is proposed. In [18], the authors propose a convolutional net architecture that can be used even when the amount of learning data is limited. A chained CNN-based age and gender classification scheme is introduced in [24], where the age classifiers are trained for different genders. The apparent age estimation task is investigated in [25]; their proposed model fuses the real value-based regression and the Gaussian label distribution based GoogLeNet; the model was tested on LAP dataset. Later, their result is improved by Antipov et al. [26] by fusing the general model and the children model.

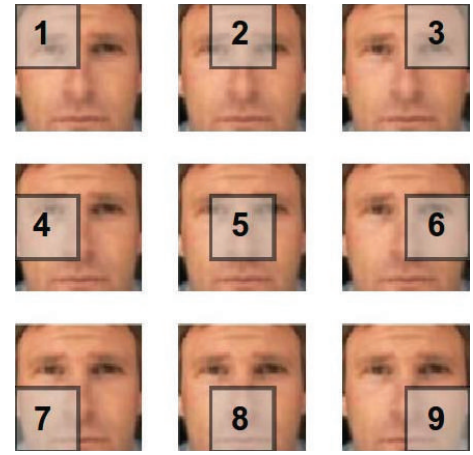


Figure 1: Illustration of the 9 image patches for local neural network training in [21].

Some researchers suggest using deeper networks for age and gender estimation. Yang et al. introduce the deep label distribution learning for apparent age estimation, where the distribution-based loss functions are used for training, which can exploit the uncertainty induced by manual labeling to learn a better model than using ages as the target [27]. The deep age distribution learning (DADL) is proposed in [28] for age prediction. Hou et al. [29] propose a deep CNN model similar with a VGG-16 net coupled with the smooth adaptive activation functions for age estimation. Their results were further improved by using the exact squared earth mover's distance in loss function [30]. In [31], convolutional neural networks are used for the extraction of deep features, then the standard support vector regression is used for gender and age prediction. Recently, the model in [31] is further improved by adding an expected value formulation after classification [32]. The directional age-primitive pattern (DAPP) is proposed in [33], which is a local face descriptor containing aging cue information; the model obtained state-of-the-art performance on Adience dataset.

2.2. Local Deep Neural Networks (LDNNs) for Age and Gender Estimation. Compare with CNN, LDNNs use a different training strategy: the small image patches around important regions of faces are extracted and used for network learning. An LDNN model for gender recognition is proposed in [20], where a feed-forward neural network without dropout is used. An edge detector is firstly used to obtain edges in face images, and small image patches are then selected around the obtained edges. All the image patches are fed into neural networks for training. The predictions of all the patches from the input test image are averaged for the final output. Using patches obtained in this way seldom leads to overfitting since the most redundant information has been removed during filtering.

Another LDNN model was proposed recently, which aims to further reduce the number of image patches used for training [21]. This model uses the same network architecture of [20]. In this modified version of LDNN, only 9 fixed image patches are used for the local network training, as presented in Figure 1. In addition, the authors split an image into

five rows (2) and find that the rows containing eye regions and mouth region are important rows for age estimation. By using less training image patches, the model still achieve a competitive performance.

3. Methodology

In this section, we describe the proposed architecture for age and gender classification. Our methodology is essentially composed of three steps: (1) to implement face detection and facial landmark localization, (2) to select image patches based on the obtained facial landmarks, (3) and to construct LDNN model. In the following, the three parts are described in detail.

3.1. Facial Landmark Localization and Patch Selection. The first step of our proposed model is to detect a face in an image and to obtain the facial landmarks on the face, both are widely investigated areas [34–36]. Currently, the global spatial models are polularly used landmark localization methods, which are mainly based on local part detectors. Therefore, it is common to use mixtures of deformable part models or to use mixtures of trees for face detection and landmark estimation. Then the efficient dynamic programming algorithms can be applied to find globally optimal solutions. Without loss of generality, a mixture of trees model for face detection and landmark localization [37] is used here. A brief introduction of the method is given below.

The model is based on a mixture of trees with a shared pool of parts V. Every facial landmark is modeled as a part; the global mixtures are used to capture topological changes due to vi a viewpoint or deformable changes such as changes in expression.

Each tree-structured pictorial structure [38] is linearly parameterized and written as $T_m = (V_m, E_m)$, where m represents a mixture and $V_m \subseteq V$. For an image I, the location of a part i is denoted as $l_i = (x_i, y_i)$. All the image parts $L = \{l_i : i \in V\}$ are scored as

$$S(I, L, m) = \mathrm{App}_m(I, L) + \mathrm{SP}_m(L) + \alpha^m, \quad (1)$$

$$\mathrm{App}_m(I, L) = \sum_{i \in V_m} \omega_i^m \cdot \phi(I, l_i), \quad (2)$$

$$\mathrm{SP}_m(L) = \sum_{ij \in E_m} a_{ij}^m dx^2 + b_{ij}^m dx + c_{ij}^m dy^2 + d_{ij}^m dy. \quad (3)$$

In (2), for the feature vector $\phi(I, l_i)$ extracted from pixel location l_i of image I, the appearance scores of placing the template ω_i^m for part i at the location l_i tuned for mixture m are summed up.

Equation (3) computes the mixture-specific spatial arrangement of parts L. $dx = x_i - x_j$ and $dy = y_i - y_j$ represent the displacement of the ith part relative to the jth part. Each term in the sum can be interpreted as a spring that introduces spatial constraints between a pair of parts [37]. The parameters $(a, b, c, \text{and } d)$ specify the rest location and rigidity of each spring.

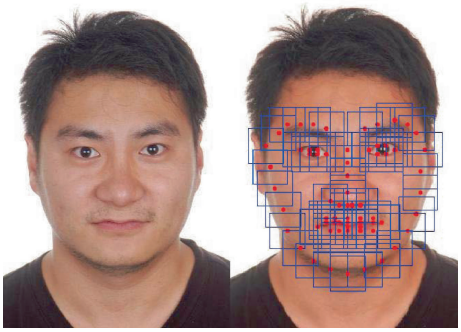


Figure 2: Landmark localization of a sample face image; the red points indicate the center of the detected landmark regions.

Figure 3: Illustration of the five rows split of a face image in [21].

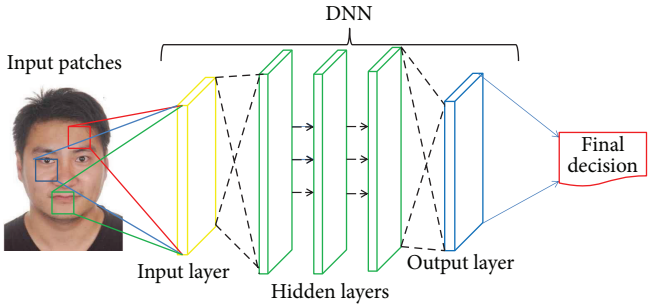


Figure 4: LDNN model used in [20, 21] is also used in this paper.

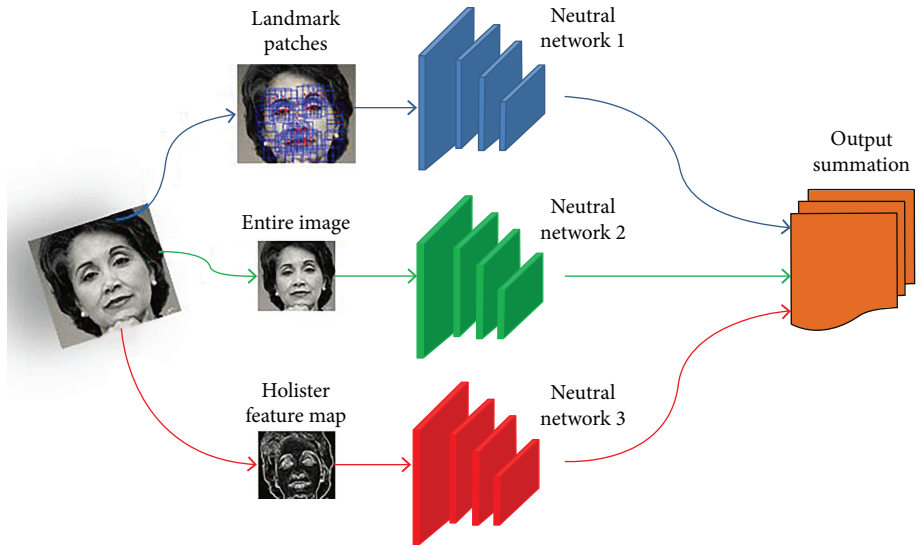


Figure 5: The proposed classification of the main architecture.

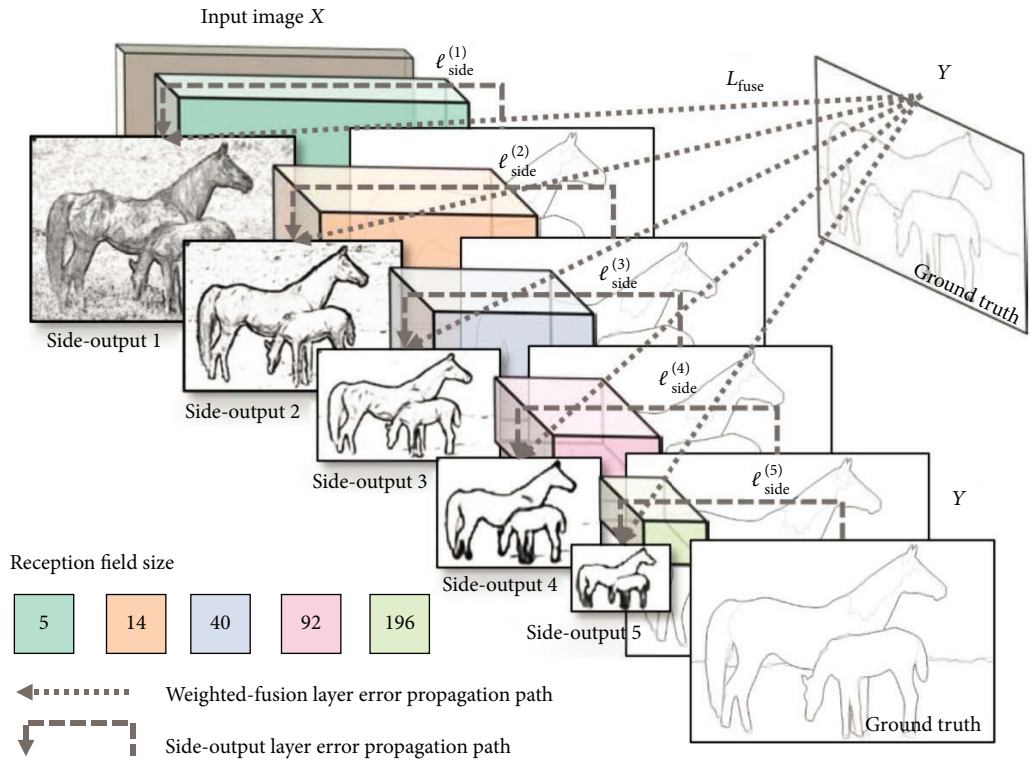


Figure 6: The HED architecture [22].

The model is trained in a fully supervised scheme, where the positive images with landmarks and mixture labels are provided, and the negative images without faces are also provided as well. The shape and appearance models are learned by using a structured predication framework. The Chow-Liu algorithm [39] is used to find the maximum likelihood tree structure which can give the best description of the landmarks in a given mixture. Figure 2 shows the landmark localization result from a sample image.

Once the facial landmarks are obtained, the local image patches can be determined. As shown in Figure 2, for the sample image, a total of 68 landmark points are detected; therefore, 68 image patches around the landmark centers are selected for the network learning. The size of the image patches in Figure 2 are 15×15 (blue squares); however, the size of an image patch can vary according to the size of the input images. In our experiments, the performances of different patch sizes are also compared.

Figure 7: Sample images from the LFW database.

The image patch selection used here is different with the two former LDNN-based methods used in [20, 21]. In [20], although the authors only keep the image patches whose center pixel is an active pixel in the binary mask image, there are still hundreds of image patches left for network training. In [21], only 9 fixed patches are used (Figure 3). However, in order to improve model performance, an image is divided into 5 rows, and the rows containing the eye regions and mouth region are used to assist the model output. The patch selection is more empirically decided in this scenario. Our landmark-based patch selection method can keep the most important information in a face image; moreover, it largely reduces the number of training patches.

3.2. Network Architecture. LDNNs are trained by using the image patches extracted from landmark regions of face images. As most of the redundant information has been discarded in our patch selection process, the left training patches cannot lead to the problem of overfitting. Therefore, it is reasonable to use a simple feed-forward neural network. The network architecture used in [20, 21] can also be directly used here for our tasks. Figure 4 shows the network architecture.

The whole procedure of our method is shown in Figure 5. For an input image, its landmark patches are detected and classified by the trained neural network. Then the outputs of the patches are averaged. Following the routine of [21], the entire image can be used to improve classification performance as well. Therefore, another neural network is trained by the entire image is also used here. Moreover, we employ the holistically-nested edge detection (HED) detector [22] to train a third neural network for further performance improvement.

The HED is a deep learning-based edge detection method; it aims to obtain a network that learns features from which it is possible to produce edge maps approaching the ground truth. HED uses multiscale and multilevel structure to generate 5 side-outputs which improve the final fusion result. The architecture of HED can be seen in Figure 6.

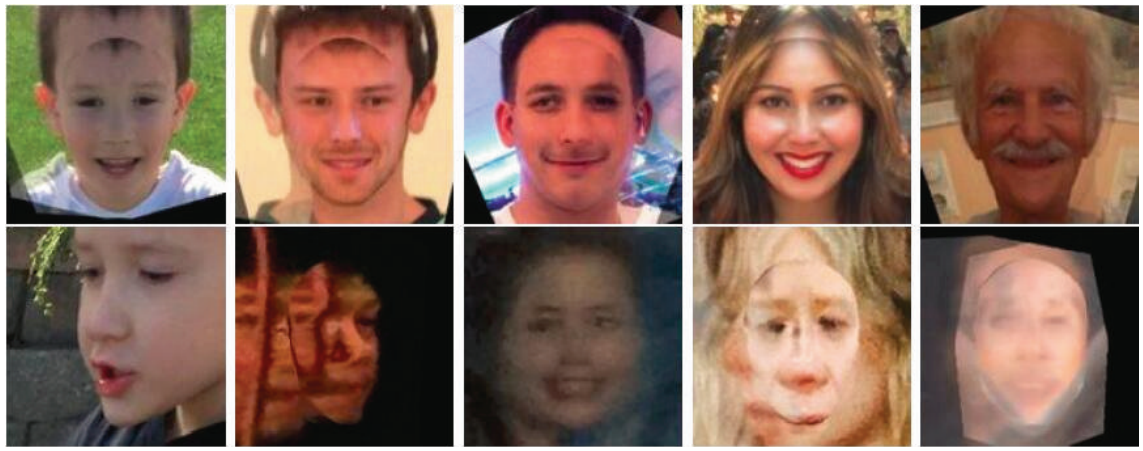


Figure 8: Sample images from Adience database.

Table 1: The label information of the Adience subset used in our experiments.

Age group	1	2	3	4	5	6	7	8	Total
Male	533	693	736	508	1635	1011	333	291	5740
Female	494	910	952	699	1867	875	296	308	6401
Total	1027	1603	1688	1207	3502	1886	629	599	12,141

In Figure 6, the side-output layers are inserted following the convolutional layers. Deep supervision is imposed at each side-output layer to guide the side-outputs toward edge predictions. The outputs of HED are multiscale and multilevel, with the side-output-plane size becoming smaller and the receptive field size is becoming larger. One weighted-fusion layer is added to automatically learn how to combine outputs from multiple scales. The entire network is trained with multiple error propagation paths (dashed lines). The details of HED can be seen in [22].

4. Experiments and Results

A series of experiments has been conducted on two popularly used face image datasets, the LFW database and the Adience database. In this section, the datasets used in our experiments are introduced firstly then the parameter settings of the experiments are introduced. Finally, the experimental results of gender and age estimation are given.

4.1. Face Image Datasets

4.1.1. Labeled Faces in the Wild (LFW). The labeled faces in the wild (LFW) database contains 13,233 face photographs labeled with the name and gender of the person pictured. Images of faces were collected from the web with the only constraint that they were detected by the Viola-Jones face detector [40]. The sample images from LFW database are shown in Figure 7.

There are four versions of LFW—the original version, funneled version, deep funneled version, and frontalized version (3D version). LFW is an imbalanced database including 10,256 images of men and 2977 images of women from 5749 subjects; 1680 of which have two or more images [40]. The 3D version is used in this work since the images are already cropped, aligned, and frontalized properly.

Table 2: The parameter settings in our experiments.

Learning algorithm	SGD + momentum
Dropout probability for input/hidden units	0.75/0.5
Initial learning rate	3
Learning rate update rule	$l_c = l_c * 0.997$ for each epoch
Initial/final momentum	0.5/0.99
Number of hidden units	512
Number of hidden layers	3
Activation function	ReLU

4.1.2. Adience Dataset. There are 26,580 face images from 2284 persons in the Adience dataset [41]. The images are with age and gender labels, which are collected from the Flickr albums and released by their authors under the Creative Commons (CC) license. The images are completely in the wild as the photos were taken under different variations in appearance, noise, pose, and lighting, and so on.

There are three versions of the Adience database, including the original version, aligned version, and frontalized version (3D version) with 26,580, 19,487, and 13,044 images, respectively. The 3D version is used in this work since most images are already frontalized and aligned to the centre of the image. However, images in the Adience database 3D version may be extremely blurry or frontalized incorrectly as shown in Figure 8. Additionally, people in the images

Table 3: The gender classification results on LFW dataset using different hidden layer numbers under the patch size 1313.

Methods compared	LDNN [20]	LDNN + locations [20]	LDNN-F [21]	Proposed	Proposed + locations
Accuracy (1 hidden layer)	91.66	92.64	94.03	94.26	94.32
Accuracy (2 hidden layers)	95.35	95.98	94.22	94.85	94.82
Accuracy (3 hidden layers)	95.81	96.04	95.64	95.53	96.02
Accuracy (4 hidden layers)	95.79	**96.25**	95.29	95.47	95.88

Table 4: Performance evaluation of different sizes of image patches.

Patch sizes	10×10	13×13	15×15	20×20	30×30
Accuracy (3 hidden layers)	95.78	**96.02**	95.63	95.86	95.54

Table 5: Classification results from different network combinations.

Network combinations	Entire image	Landmark patches	Entire image + landmark patches	Combined
Classification accuracy	92.86	95.06	95.87	**96.02**

could show emotions. Therefore, it is reported that patches extracted from those images may not always contain the same face region which may result in lower classification rates [21].

There are three subsets of the Adience dataset 3D version; this is because it is not necessary to label gender with age groups or vice versa. The first subset contains 12,194 images labeled with gender. The second subset comprises 12,991 images labeled with age. 12,141 images are included in the third subset, which is labeled as both gender and age. Our experiments are run on the third subset. The label information can be seen in Table 1.

4.2. Experimental Settings. In order to find appropriate parameters for the proposed method, a series of experiments has been conducted. The parameters listed in Table 2 produced good outcomes.

The experiments were run on a PC with an Intel i7 4 cores CPU, 16 G memory and an NVIDIA Geforce GTX 1080 GPU (8 G memory); the time cost for training the proposed model is around 10 hours.

4.3. Experimental Results on LFW. For comparison, we follow the routines in [20, 21] to carry out our experiments. Five cross-validations using the same five folds as [20, 21] are used. Around 67% of patches of men are randomly discarded in each fold to balance the data. We first set the size of the image patches as 13×13, which is the same as described in [20]. Table 3 lists the classification results of our method and the compared methods, where different numbers of hidden layers are also tested.

In Table 3, one can see that for the same model, besides the image patches themselves, if the center locations' coordinates are added to indicate where a patch is extracted (the "LDNN + location" and "proposed + location" columns), the classification performance can be improved. The method in [21] uses 9 fixed image patches; therefore, the location of patches are also fixed; the method is named as LDNN-F in Table 3.

Table 6: Gender classification results on LFW dataset from different methods.

Methods compared	Accuracy (%)
LDNN	96.25
LDNN-F	95.64
Compact CNN [26]	**97.03**
LBP + SVM [10]	95.6
Gabor + PCA + SVM [42]	94.01
Proposed	96.02

Table 7: Gender classification results on the Adience dataset.

Compared methods	Entire image	LDNN-F [21]	Proposed
Classification accuracy	77.84	78.63	**80.64**

In Table 3, the best gender classification is 96.25% from [20] with patch location. Due to the huge number of patches and the limited amount of memory, it is not feasible to train a neural network using all of the four training folds. In the same way in [21], only one fold was used for network training in this work. The smaller training set is a factor leads to a lower performance.

The effect of using different sizes of image patches are also evaluated. Three hidden layers are used in the network. The compared results can be seen in Table 4. The best performance among the tested patches sizes was obtained by 13×13, which is the same with the results in [20], where the authors explain the size of 13×13 was determined from a previous research. The best performance of [21] was obtained by using a larger size as 30×30. The reason is only 9 location-fixed patches are used for training; a larger patch is able to contain more useful information.

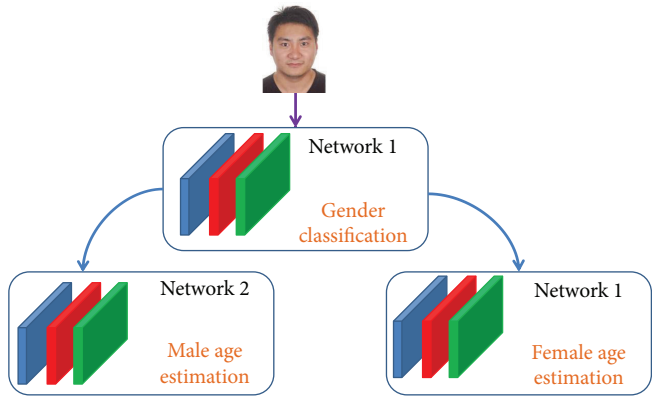


Figure 9: The age estimation scheme.

In our proposed model, besides the landmark-based image patches, the entire image and the holistic feature map extracted from the entire image are also used to further improve model performance. Table 5 lists the results of improvement bought by the holistic feature.

Some of the state-of-the-art methods work on the LFW dataset for gender classification are also compared in our experiments. The results are listed in Table 6. Among the compared methods, the best performance 97.03% was obtained by the method of "Compact CNN"; however, this method needs to construct an ensemble of learning models, which is much more complicated on model construction compared to our method.

4.4. Experimental Results on the Adience Dataset. The age and gender classification are run on the 3D version of the Adience dataset. We used the same routine in [21]; the networks are first trained separately for age and gender then the gender classification results are used to help age estimation.

The same parameters listed in Table 2 are also used here for model training. The performance of our model is shown in Table 7. Our proposed model achieves 80.64% correction rate on the data set, where the result in [21] is 78.63%. The main reason is the Adience dataset are not frontalized well; the location-fixed patches used in [21] may not always contain the same region of faces. In our method, by detecting facial landmarks in advance, the obtained landmark-based patches can relieve this problem much better.

The Aidence dataset contains 8 age groups and another 20 different age labels. Some folds even lack the images for some age groups; therefore, the age labels must be merged. We used the same merging scheme used in [21]; all the labels are merged into the 8 age groups. Please see their paper for details.

In the same way in [18, 21], the one-off classification rate is used for age estimation. That is due to the apparent similarity of persons in adjacent age groups; images which are categorized into adjacent age groups are considered to be correct classification.

For the age estimation, three sets of neural networks are constructed; each contains the model shown in Figure 3. The neural network 1 is used for gender classification, and neural network 2 and 3 are for male and female age estimation, respectively. If an input image is recognized as male then network 2 will be used for its age estimation; otherwise, network 3 will be activated. The whole process can be seen in Figure 9.

Table 8: Age estimation results from the two neural networks for men and women respectively.

Rate method	Neural network 1 (male's age)		Neural network 2 (female's age)	
	Exact	One-off	Exact	One-off
Entire image	38.94	77.76	36.68	75.12
LDNN-F [21]	39.90	80.32	41.27	77.14
Proposed	41.86	81.87	42.79	78.65

Table 9: Age estimation results from the proposed age estimation model compared with other CNN-based methods.

Methods	LDNN-F		CNN [41]		Proposed	
	Exact	One-off	Exact	One-off	Exact	One-off
Accuracy	41.82	77.98	45.1 ± 2.6	79.5 ± 1.4	**44.36**	**80.69**

It should be noted that the neural network 2 in Figure 9 is trained using 5740 face images of men, and the network 3 is trained using 6410 images of women. The individual performance of neural network 1 and neural network 2 on age estimation is shown in Table 8, and the results from the model in Figure 9 is given in Table 9.

5. Conclusion and Future Work

A modified version of local deep neural networks is proposed in this paper. Instead of using location-fixed patches, the facial landmarks are detected in advance, then the image patches around landmarks are selected for network training, which greatly reduces the training cost. Moreover, the experimental results show that the method proposed in this paper achieves competitive performance in the two tested datasets. The performance of the proposed model still can be improved by incorporating other schemes into current architecture, for example, to use a more efficient facial landmark detection method or to further optimize the network structure, these will be investigated in our future work.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Additional Points

The ownership of Figures 2, 4, and 9 in this paper belongs to the original author Yungang Zhang. Please do not reprint, duplicate, or use these pictures in any form without the permission from the author. Otherwise, the author will have the right to investigate for legal liability.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The project is funded by Natural Science Foundation China Grant no. 61462097 and Yunnan Provincial Education Department research Grant no. 2018JS143.

References

[1] B. A. Golomb, D. T. Lawrence, and T. J. Sejnowski, "Sexnet: a neural network identifies sex from human faces," in *Advances in Neural Information Processing Systems*, vol. 3, pp. 572–577, Denver, CO, USA, 1991.

[2] S. Tamura, H. Kawai, and H. Mitsumoto, "Male/female identification from 8×6 very low resolution face images by neural network," *Pattern Recognition*, vol. 29, no. 2, pp. 331–335, 1996.

[3] B. Moghaddam and Ming-Hsuan Yang, "Learning gender with support faces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 707–711, 2002.

[4] J. Yang, D. Zhang, A. F. Frangi, and J.-Y. Yang, "Two-dimensional PCA: a new approach to appearance-based face representation and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 1, pp. 131–137, 2004.

[5] C. Shan, "Learning local binary patterns for gender classification on real-world face images," *Pattern Recognition Letters*, vol. 33, no. 4, pp. 431–437, 2012.

[6] J. G. Wang, J. Li, W. Y. Yau, and E. Sung, "Boosting dense sift descriptors and shape contexts of face images for gender recognition," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*, pp. 96–102, San Francisco, CA, USA, 2010.

[7] X. M. Leng and Y. D. Wang, "Improving generalization for gender classification," in *2008 15th IEEE International Conference on Image Processing*, pp. 1656–1659, San Diego, CA, USA, 2008.

[8] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, pp. 886–893, San Diego, CA, USA, 2005.

[9] G. Guo, G. Mu, Y. Fu, and T. S. Huang, "Human age estimation using bio-inspired features," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 112–119, Miami, FL, USA, 2009.

[10] J. E. Tapia and C. E. Perez, "Gender classification based on fusion of different spatial scale features selected by mutual information from histogram of LBP, intensity, and shape," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 3, pp. 488–499, 2013.

[11] H. Han, C. Otto, and A. K. Jain, "Age estimation from face images: human vs. machine performance," in *2013 International Conference on Biometrics (ICB)*, pp. 1–8, Madrid, Spain, 2013.

[12] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional networks," in *Advances in Neural Information Processing Systems 25 (NIPS 2012)*, pp. 1097–1105, Lake Tahoe, NV, USA, 2012.

[13] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun, "Overfeat: integrated recognition, localization and detection using convolutional networks," 2013, https://arxiv.org/abs/1312.6229.

[14] A. Romero, N. Ballas, S. E. Kahou, A. Chassang, C. Gatta, and Y. Bengio, "Fitnets: hints for thin deep nets," 2014, https://arxiv.org/abs/1412.6550.

[15] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, https://arxiv.org/abs/1409.1556.

[16] J. Gu, Z. Wang, J. Kuen et al., "Recent advances in convolutional neural networks," *Pattern Recognition*, vol. 77, pp. 354–377, 2018.

[17] S. Chen, C. Zhang, M. Dong, J. Le, and M. Rao, "Using ranking-cnn for age estimation," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 742–751, Honolulu, HI, USA, 2017.

[18] G. Levi and T. Hassncer, "Age and gender classification using convolutional neural networks," in *2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 34–42, Boston, MA, USA, 2015.

[19] J. van de Wolfshaar, M. F. Karaaba, and M. A. Wiering, "Deep convolutional neural networks and support vector machines for gender recognition," in *2015 IEEE Symposium Series on Computational Intelligence*, pp. 188–195, Cape Town, South Africa, 2015.

[20] J. Mansanet, A. Albiol, and R. Paredes, "Local deep neural networks for gender recognition," *Pattern Recognition Letters*, vol. 70, pp. 80–86, 2016.

[21] Z. Liao, S. Petridis, and M. Pantic, "Local deep neural networks for age and gender classification," 2017, https://arxiv.org/abs/1703.08497.

[22] S. Xie and Z. Tu, "Holistically-nested edge detection," in *2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 1395–1403, Santiago, Chile, 2015.

[23] D. Yi, Z. Lei, and S. Z. Li, "Age estimation by multi-scale convolutional network," in *Asian Conference on Computer Vision*, pp. 144–158, Singapore, 2014.

[24] A. Ekmekji, "Convolutional neural networks for age and gender classification," Technical Report, Stanford University, 2016.

[25] X. Liu, S. Li, M. Kan et al., "Agenet: deeply learned regressor and classifier for robust apparent age estimation," in *2015 IEEE International Conference on Computer Vision Workshop (ICCVW)*, pp. 258–266, Santiago, Chile, 2015.

[26] G. Antipov, M. Baccouche, S.-A. Berrani, and J.-L. Dugelay, "Apparent age estimation from face images combining

general and children-specialized deep learning models," in *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 96–104, Las Vegas, NV, USA, 2016.

[27] X. Yang, B.-B. Gao, C. Xing et al., "Deep label distribution learning for apparent age estimation," in *2015 IEEE International Conference on Computer Vision Workshop (ICCVW)*, pp. 344–350, Santiago, Chile, 2015.

[28] Z. Huo, X. Yang, C. Xing et al., "Deep age distribution learning for apparent age estimation," in *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 17–24, Las Vegas, NV, USA, 2016.

[29] L. Hou, D. Samaras, T. M. Kurc, Y. Gao, and J. H. Saltz, "Neural networks with smooth adaptive activation functions for regression," 2016, https://arxiv.org/abs/1608.06557.

[30] L. Hou, C. P. Yu, and D. Samaras, "Squared earth mover's distance-based loss for training deep neural networks," 2016, https://arxiv.org/abs/1611.05916.

[31] R. Rothe, R. Timofte, and L. Van Gool, "Some like it hot - visual guidance for preference prediction," 2015, https://arxiv.org/abs/1510.07867.

[32] R. Rothe, R. Timofte, and L. Van Gool, "Deep expectation of real and apparent age from a single image without facial landmarks," *International Journal of Computer Vision*, vol. 126, no. 2–4, pp. 144–157, 2018.

[33] M. T. B. Iqbal, M. Shoyaib, B. Ryu, M. Abdullah-Al-Wadud, and O. Chae, "Directional age-primitive pattern (DAPP) for human age group recognition and age estimation," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 11, pp. 2505–2517, 2017.

[34] M.-H. Yang, D. J. Kriegman, and N. Ahuja, "Detecting faces in images: a survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 1, pp. 34–58, 2002.

[35] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: a literature survey," *ACM Computing Surveys*, vol. 35, no. 4, pp. 399–458, 2003.

[36] Y. Sun, X. Wang, and X. Tang, "Deep convolutional network cascade for facial point detection," in *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3476–3483, Portland, OR, USA, 2013.

[37] X. Zhu and D. Ramanan, "Face detection, pose estimation, and landmark localization in the wild," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2879–2886, Providence, RI, USA, 2012.

[38] Y. Yang and D. Ramanan, "Articulated pose estimation with flexible mixtures-of-parts," in *CVPR 2011*, pp. 1385–1392, Colorado Springs, CO, USA, 2011.

[39] C. Chow and C. Liu, "Approximating discrete probability distributions with dependence trees," *IEEE Transactions on Information Theory*, vol. 14, no. 3, pp. 462–467, 1968.

[40] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: a database for studying face recognition in unconstrained environments," Technial Report 07-49, University of Massachusetts, Amherst, 2007.

[41] E. Eidinger, R. Enbar, and T. Hassner, "Age and gender estimation of unfiltered faces," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 12, pp. 2170–2179, 2014.

[42] P. Dago-Casas, D. Gonzalez-Jimnez, L. L. Yu, and J. L. Alba-Castro, "Single- and cross- database benchmarks for gender classification under unconstrained settings," in *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, pp. 2152–2159, Barcelona, Spain, 2011.

Hindawi
Journal of Sensors
Volume 2018, Article ID 4390703, 10 pages
https://doi.org/10.1155/2018/4390703

Research Article

Video Retrieval System for Meniscal Surgery to Improve Health Care Services

Sana Amanat,[1] Muhammad Idrees,[2] Muhammad Usman Ghani Khan,[1] Zahoor Rehman,[3] Hangbae Chang,[4] Irfan Mehmood,[5] and Sung Wook Baik[5]

[1]*Department of Computer Science and Engineering, University of Engineering and Technology Lahore, Lahore, Pakistan*
[2]*Department of Computer Science and Engineering, University of Engineering and Technology Lahore Narowal Campus, Lahore, Pakistan*
[3]*Department of Computer Science, Comsats Institute of Information Technology Attock Campus, Attock, Pakistan*
[4]*Department of Industrial Security, College of Business and Economics, Chung-Ang University, Seoul, Republic of Korea*
[5]*Department of Software, Sejong University, Seoul, Republic of Korea*

Correspondence should be addressed to Irfan Mehmood; irfanmehmood@ieee.org

Received 18 January 2018; Accepted 29 April 2018; Published 14 June 2018

Academic Editor: Ka L. Man

Meniscal surgery is considered the most general orthopedic process that deals with the treatment of meniscus tears for human health care. It leads to a communal contusion to the cartilage that stabilizes and cushions the knee joints of human beings. Such tears can be classified into different categories based on age group, region, and occupation. Further, a large number of sportsmen and heavy weightlifters even in developed countries are affected by meniscus injuries. These patients are subjected to arthroscopic surgery, and during surgical treatment, the perseverance of meniscus is a very crucial task. Current research provides a significant ratio of meniscal tear patients around the globe, the critical expanse is considered as having strikingly risen with a mean annual of 0.066% due to surgery failure. To decumbent this ratio, an innovative training mechanism is proposed through video retrieval system in this research. This research work is focussed on developing a corpus and video retrieval system for meniscus surgery. Using the proposed system, surgeons can access guidance by watching the videos of surgeries performed by an expert and their seniors. The proposed system is comprised of four approaches to the spatiotemporal methodology to improve health care services. It entails key point, statistical modeling, PCA-scale invariant feature transform (SIFT), and PCA-Gaussian mixture model (GMM) with a combination of sparse-optical flow. The real meniscal surgery dataset is used for testing purposes and evaluation. The results conclude that using PCA-SIFT approach improves the results with an average precision of 0.78.

1. Introduction

A meniscus tear is a public discolouration of the cartilage that alleviates and pads the knee joints of human beings. This is common malfunctioning across the globe, with a mean annual rate of 0.066%. It is noteworthy to mention that a large number of meniscus injury incidences are observed at about 66 per 100,000 around the globe. These incidences can be categorized under various aspects, that is, age group, region, and occupation. Most of the sportsmen and heavy weightlifters affected with age less than 40 years in developed countries recovered using meniscus surgery. They cover the areas of America, Latin, Asia, and Africa. On the other hand in Denmark, more degenerative changes in the annual incidence are observed. These changes are increasing from 164 to 312 over the past 10 years and about 2 million cases observed in meniscus surgery treatment. Details of registered patients, diagnosis, and procedures maintained by Danish National Patient Register can be found in [1] and Summit (2016).

Previous studies have reported a very high rate of repetitive surgery due to the failure of meniscus surgery. Different causes are observed for the failure of meniscus surgery. One of the major causes is the inappropriate removal of meniscus

because unexperienced surgeons do not know how much or which part has to be removed properly. More precisely, it deals with the interaction of tissues and radiation, which may be represented in the form of a static image and video or recordings. These recordings can be used to train surgeons and practitioners to perform successful surgeries. One way to access these recordings is through video retrieval system that aims to retrieve videos from a source of the multimedia repository. Upon surgeon's request and need, video retrieval system can provide the most relevant and accurate videos of meniscal surgery. Likewise, there have been several studies in relation to surgery, that is, for the eyes, lungs, heart, and kidney, but meniscal surgery holds a unique position since it is a commonality in most of the countries and in literature neither corpus nor video retrieval systems precisely for meniscal surgery yet found [2].

To retrieve retinal surgery videos using compressed video streams, the required motion info is extracted at regular intervals. Precisely, two types of motion information are excerpted: one is region trajectory and the other encompasses the residual error. A heterogeneous feature vector is used to combine the motion information, and sequence segmentation is utilized to extract the motion information. The Kalman filter and generalized Gaussian distribution are utilized to formulate region trajectory and residual information. In last step, extension of fast dynamic time warping is also used along with feature vector to compare videos [3].

Another approach to retrieve medical videos is a key frame which is initially extracted and its global features are excerpted, which has to be saved in XML script format in the form of an object to reduce the storage space. Furthermore, neural network and naïve Bayes are used to enhance the performance. Features of an image are extracted using entropy, Huffman coding, and fast Fourier transforms [4].

In order to retrieve laparoscopy videos from a colossal repository of videos, a feature signature is utilized which is composed of abstract and local details of the image. Initially, position, colour, and texture features of an image are considered for this signature. Weight is assigned to each feature describing the potential to be used as a feature signature representative. For comparison between two signatures, earth mover's distance, signature quadratic form distance, and signature matching distance are used in combination [5]. To retrieve endoscopy videos from repository, feature signature is computed based on global features, that is, colour, position, and texture. Weight is assigned to each feature, and zero-weighted features are considered as irrelevant for signature. Those features have weighted value unequal to zero denoted as signature representatives. Furthermore, cluster centroid of these features is used as feature representative. To measure the similarity between endoscopic images, signature matching distance and adaptive binning feature signature are used by using the visual characteristics [6].

In another approach to retrieve endoscopic videos, local, global, and feature fusion techniques are employed. To extract features, various algorithms CEDD, SURF, and SIMPLE are used. Additionally, early and late fusion approaches are also considered. These approaches are tested on 1276 video clips. LIRE software library is used to index test frames also capable to extract 20 visual features of an image [7].

To retrieve cataract surgery videos from the repository, a video is divided into a number of overlapping clips. Each clip contains the same number of frames. Various features are extracted from these frames using histogram of oriented gradients (HOG), BoVW, and motion histogram. These features are statistically modeled using the Bayesian network, CRFs, and HMMs. Thirty cataract surgery videos are used for the testing purpose having 720×576 pixels. A mean area under ROC of 0.61 was achieved in recognizing surgical steps from video stream [8]. In order to retrieve and analyse female reproductive system through hysteroscopy, a video summarization framework is proposed. To compute feature set, colour information is extracted from key frames using COC colour model. Other salient features are also extracted such as texture, motion, curvature, and contrast [9].

Another approach to retrieve cataract videos from the source that formulate a feature set is by using global features such as colour, texture, and motion information of the frame. Furthermore, the wavelet transform is performed on these extracted channels of colour. In this study, cataract and epiretinal surgeries are considered [10]. To retrieve similar videos in the domain of retinal surgery, a motion-based approach has been found in literature which utilized MPEG-4 to extract features from the frame. For video encoding, a group of pictures and macroblocks are used where three types of frames are considered. Motion histogram, classification, and motion trajectories are used to formulate motion feature set. Furthermore, signatures are also composed and to measure their ratio of similarity, fast dynamic time warping approach is used. This approach is applied to epiretinal membrane surgery.

In order to control the failure rate of surgery and to overcome the mentioned drawbacks, there is a dire need for a sophisticated approach. It could be possible via the fabrication of a connexion between information technology and medical sciences. Information technology is capable of providing medical diagnosis, clinical analysis, and treatment planning through a practice known as medical imaging [11, 2, 12].

Our proposed system firstly focuses on spatial features, that is, the location of objects in reference to static objects [13]. Secondly, it deals with the temporal information of an image, which provides a glimpse of moving and operational objects [14]. This system is comprised mainly of four methods, that is, interest point, statistical modeling, PCA-SIFT, and PCA-GMM-based video retrieval. The first method incorporates interest points of videos which are excerpted using up-to-the-minute feature explication and indicator algorithms known as SIFT, HOG, and covariant feature detector (CFD) with distance transform (DT) addressed by Bai et al. [15] and Bharathi et al. [16]. The second method encompasses statistical modeling of features using GMM and Kullback-Leibler divergence. Moreover, in the third and fourth method, dimension reduction technique is integrated with SIFT and GMM, respectively. To extract the temporal features, Lucas-Kanade sparse-optical flow is incorporated in each method [17]. This proposed system can help new

and less-experienced surgeons to get appropriate information about surgical methods, procedures, and tools to perform a successful surgery [11, 2, 12]. These meniscal recordings can additionally be used for the explanation of the patients and their follow-up sessions [9].

Table 1 illustrates a summarized comparison of previous approaches to retrieve videos. It is clear from this comparison that there are so many systems for medical video retrieval, but a video retrieval system for meniscal surgery practitioners is not found in spite of its need around the globe. In this system, we also have used a hybrid of state-of-the-art techniques for feature detection, recognition, and comparison. The most important point that can be concluded in this summarization is a step towards new research direction.

2. Materials and Methods

This section sheds light on the proposed methodology for meniscal surgery video retrieval. The surgical video is considered to be composed of frames which are chronologically bound together to define a tale, that is, live surgical process. Each frame grasps an instant of surgery, producing a coherent association with contiguous frames. Algorithms in this framework are applied to the frame instead of a video because the frame is a building block of a video. Therefore, these videos are decomposed into frames and shot boundary is detected using XOR. As a result, the collections of master frames depicting master shot are extracted. These master frames are further utilized to model spatiotemporal features of a particular video. Detail of spatiotemporal feature extraction and modeling technique is as follows.

2.1. Spatiotemporal Modeling Using Interest Points. The interest point represents the identification of interest or key points within an image used for further processing. A description of the interest point is as follows:

(i) It has a demarcated location and mathematical description of an image.

(ii) Local image organization is high in the context of local information content.

(iii) In the image domain, it has steady trepidation of both local and global.

2.2. SIFT and CFD-Based Approach. SIFT is used to detect interest points using various scales. With the help of Gaussian filters, the change in succeeding Gaussian-blurred images is figured out to excerpt interest points [17]. A difference of Gaussian (DoG) image $\mathrm{DoG}(a, b, \Delta)$ of an original image Im (a, b) is given by

$$\mathrm{DoG}(a, b, \Delta) = L(a, b, L_i\Delta) - L\left(a, b, L_j\Delta\right), \tag{1}$$

where $L(a, b, L\Delta)$ describes the convolution operation of the original image Im (a, b) with Gaussian blur $G(a, b, L\Delta)$, at scale $L\Delta$. [19].

CFD uses an affine adaptation to approximate the figure of an affine region in order to compute covariant features of an image. With the help of affine adaptation, interest points can be calculated as follows [10]:

$$\mathrm{Cfd}(U) = \begin{bmatrix} \mathrm{De}_{aa}(U) & \mathrm{De}_{ab}(U) \\ \mathrm{De}_{ab}(U) & \mathrm{De}_{bb}(U) \end{bmatrix}, \tag{2}$$

where $\mathrm{De}_{aa}(U)$ is defined as the 2nd fractional derivative in the direction a and $\mathrm{De}_{ab}(U)$ represents the mix fractional 2nd derivative in directions of a and b. These computed features of input and target images are matched using Euclidean distance to get the value of similarity [20]. These calculated distances are sorted to make the ranked list based on similarity of input video with target videos.

2.3. HOG and Distance Transformation (DT). The HOG feature detector uses edge directions to compute local object appearance and shape within an object.

Each image is split into small cells, and histogram of gradient directions is computed for all cells.

In order to improve accuracy, contrast normalization of local histograms is computed by measuring the intensity across image block. Dalal-Triggs method can be used for normalization such as [3].

$$f = \frac{\mathbf{N}}{\sqrt{\|N\|_2^2 d^2}}. \tag{3}$$

All histograms in a given block are represented by a vector $\mathbf{N}$, which is not normalized, a small constant value is represented by d and $\|N\|k$ be its k-norm for $k = 1, 2$ [21].

DT [3] descriptor is also called distance map which labels the pixels by measuring distance with the support of Canny's edge detector. A framework of key point base is present in Figure 1.

2.4. Statistics-Based Framework for Video Retrieval. Statistics based technique is an approach to scrutinize condense and to interpret data. For this purpose, we have used GMM with Kullback-Leibler divergence to model the spatiotemporal content of videos. Following is the description of a method.

2.5. Spatiotemporal Modeling through GMM. The surgical footage is an object which contains highly complex multidimensional contents based on time series folded data. Assume that a video V is an arrangement of frames and has length l, $\{v_1, \ldots, v_n\}$, with every frame v_i, providing a sequential pattern of an arrangement.

A video frame is described by each element v_i, and its spatial features are modeled in d-dimensional space. At this phase, i_s presents the spatial features of frame x_i in d dimension feature space. Spatial features of a frame constitute facts of locality with the magnitude of features, for example, colours of every pixel (red, green, and blue), and collected objects may be easily traced. In order to model RGB colour distribution, pixel position can be stated as shown in Figure 2.

A spatiotemporal probabilistic model can be constructed through statistical observation of samples by using time and space features. The sampling of the video is performed at a

Table 1: Comparison of proposed methodology with existing ones.

Reference papers	Feature set	Feature descriptor	Similarity/cluster/ classification measurement	Dataset	Research domain	Video type
Droueche et al. [18]	A motion vector, residual information	Kalman filter, generalized Gaussian distribution	Extension of fast dynamic time warping (EFDTW)	69 video-recorded retinal surgery steps, 1400 video-recorded cataract surgery, and 1707 movie clips with classified human actions	Eye surgery	MPEG-4 AVC/H.264
Ramya et al. [4]	Global features	Entropy, histogram, frequency	Neural network	15478 XML scripts based on key frames	Medical videos	mp4
Schoeffmann et al. [5]	Position, colour, texture	N/A	Earth mover's distance, signature quadratic form distance, and signature matching distance	25 stored video files per procedure	Laparoscopic surgery	N/A
Beecks et al. [6]	Spatial information (colour, coarseness, contrast value)	Local feature descriptor	*k*-mean cluster	25 segment videos per procedure, resulting in total 1276 short video clips ranging from 60 sec to 250 sec	Endoscopic videos	N/A
Carlos et al. [7]	Local and global features, feature fusion	SIMPLE, CEDD, SURF	*k*-mean	1276	Endoscopy videos	N/A
Charriere et al. [8]	N/A	BoVW	*k*-nearest neighbour (KNN), motion histogram, and hidden Markov model	30	Cataract surgery	DV format
Muhammad et al. [9]	Colour opponent colour space (CoC), curvature map, texture	N/A	N/A	10	Hysteroscopy	N/A
Quellec et al. [10]	Texture, colour, optical flow	N/A	N/A	23 retinal, 100 cataract surgery, 69 hollywood movie	Eye surgery	MPEG 2
Droueche et al. [18]	Motion histogram	Fast dynamic time warping	*k*-mean, Kalman filter, and extended of fast dynamic time warping	23 epiretinal surgery video	Epiretinal membrane surgery	MPEG-4 AVC/H.264
Proposed	The key point, colour	SIFT, HOG, DT, CFD, PCA-SIFT, spatiotemporal feature vector	GMM, Euclidean distance, and Kullback-Leibler divergence	20 meniscus surgery	Meniscus surgery	Flv, Avi, mp4

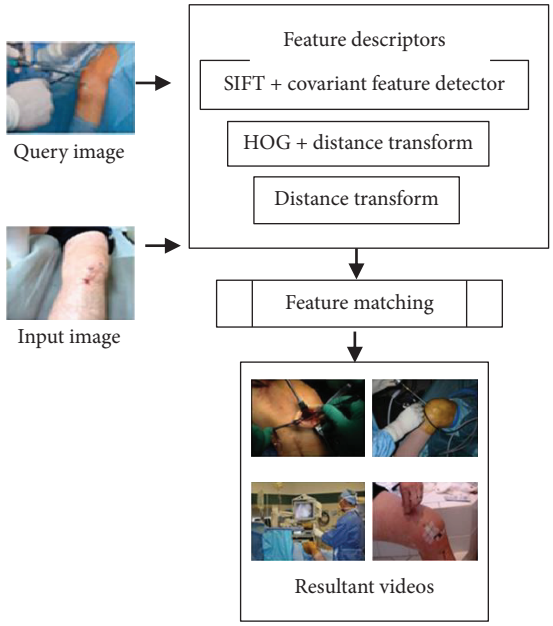


Figure 1: Framework of key-point-based retrieval.

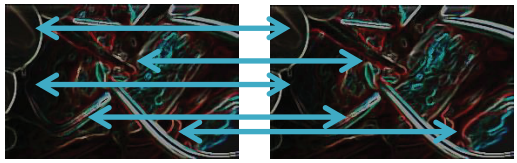


Figure 2: Temporal connection is created to form continuity between the consecutive frames.

presumed rate, preserving temporal enslavement by their adjoining images. Spatial features of every image depend on the spatial context of every image and may depend on its former image, which is known as spatiotemporal association.

Furthermore, GMM [22] is applied to an approximate likelihood density of collected items. GMM is an unsupervised learning approach in which probabilistic model is built with the assumption that data points are produced by an assortment of a finite amount of Gaussian distributions through unidentified parameters.

Moreover, a d-dimensional spatial element for the image x_i may be demonstrated through GMM along with K quantities and a presumed number of a frame structure is represented by K, for instance, the numeral core objects instituted in particular section. By using MLE, an association can be computed through the probability of spatiotemporal continuousness amongst two successive images as presented in Figure 3.

We have acquired Gaussians against input and output videos after applying maximum likelihood as shown in Figures 4 and 5. By analyzing data distribution of these Gaussians, we can determine that dissimilarity is higher between input and output. To measure the dissimilarity ratio, k-Lealer divergence, also known as relative entropy, was used.

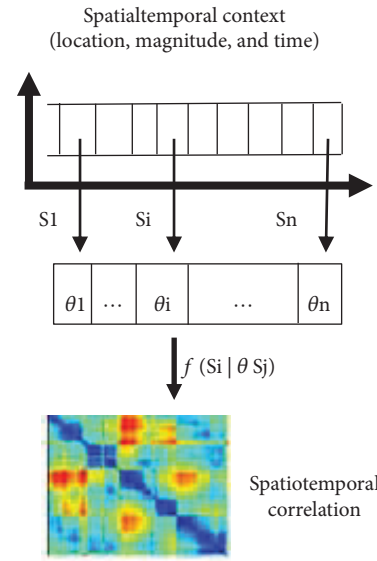


Figure 3: Spatiotemporal probabilistic model for a video frame sequence $\{x1, \ldots, xn\}$.

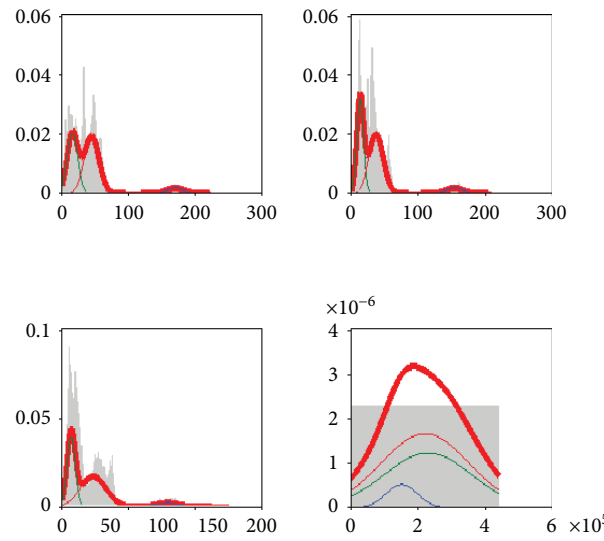


Figure 4: GMM of an input query.

2.6. Compression-Based Video Retrieval. In order to reduce insignificance and redundancy, datum compression technique can be used. The intention of compression is to cut down the data dimensions for effective storage, retrieval, and manipulation. For this purpose, principal components analysis (PCA) is used to the significant components from data that immaculately defines the entire image. Moreover, a hybrid approach is applied by the combination of PCA along with SOFT and GMM to improve the precision of retrieval method. Details of the method are as follows.

2.6.1. Summary of PCA. Principal component analysis is the dimensional reduction method which calculates linear projections to maintain maximum variance in original data.

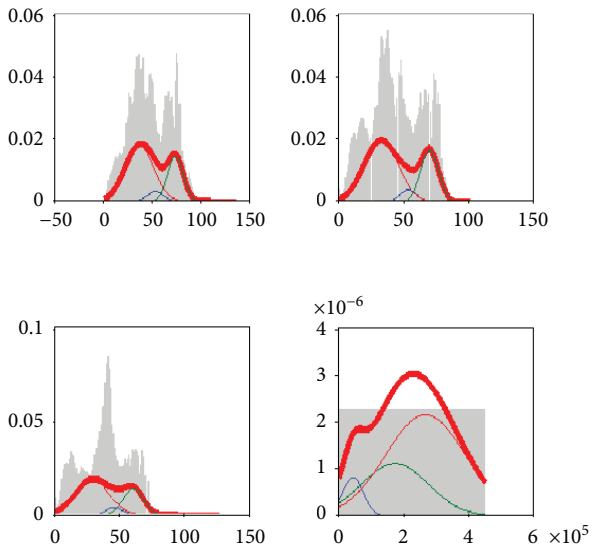


Figure 5: GMM of the output query.

Eigenvectors to Eigenvalues $\boldsymbol{\lambda}$ of a matrix of the covariance E_x of datum $D \in R^{dxM}$ where d defines the dimensions of the original data and M shows the total count of points.

$$E_x a = \lambda a. \tag{4}$$

By visualizing the point x_i via projecting

$$y_i = Ax_{\text{is}}. \tag{5}$$

A defines a matrix having eigenvectors associated with 2 or 3 of the biggest eigenvalues, and Y_i is a calculated lower dimension sketch of X_i. PCA is the most commonly used method to visualize new data. Images produced by PCA are very easy to deduce and also works moderately well even when the difference in data stands primarily intense in solitary limited directions [23].

2.6.2. PCA-SIFT-Based Video Retrieval. In this method, PCA is applied on SIFT which is highly vigorous to image distortion. In PCA-SIFT firstly computes an Eigenspace to define an inline image of the native patch. In the second step, the local image gradient is calculated on a given patch. In the third step, it projects gradient image vector by the Eigenspace to originate the dense feature trajectory. That feature trajectory is expressively negligible than SIFT feature trajectory and may be used with the identical matching method. Euclidean distances amongst binary feature trajectories can be used to define either these two trajectories resemble identical interest point in dissimilar frames. The following are the steps to retrieve videos using PCA-SIFT approach.

(1) Input shot is transformed to image.

(2) The master image is extracted using shot boundary detection.

(3) For each mater frame, PCA-SIFT is calculated.

(4) Steps from 1–3 are repeated for target videos from 1–N.

(5) After step 4, two feature vectors were gotten, that is, for input and target videos [1 – N].

(6) The feature vector of input and target videos is compared using Euclidean distance to compute the similarity.

2.6.3. PCA-GMM-Based Video Retrieval. The analyses of Figures 4 and 5 show that spatial context is comprised of location and magnitude features against each pixel such as RGB colour channel. These colour channels are extracted from input and target videos and passed through dimension reduction process using PCA. These compressed channels are modeled using GMM for spatiotemporal analysis. Furthermore, Kullback-Leibler divergence is applied in similar way as presented in Section 3.2.

2.6.4. Temporal Points Extraction from Surgical Videos by Sparse-Optical Flow. Surgical video is a dynamic entity in which the motion of an object plays a vital role in understanding an ongoing surgical process. To capture the movement or motion of surgical procedure, we have used optic flow method, which is an arrangement of moving object in the photographic sight prompted through a relative movement amongst a scene plus an eye, that is, camera. The Lucas-Kanade method was used to capture the motion of objects within a video of live surgery [17].

3. Results and Discussion

In this section, we present experimental settings including the data set and evaluation method.

3.1. Dataset. Datasets are comprised of 20 real meniscus surgical videos collected from various online publicly available repositories MedlinePlus [24]; AAOS [25]; vjortho (2017); and Northgate (2017). These shots are sufficient and beneficial in the current scenario because it covers a wide range of multiple scenarios and uniquely identifies a number of instances. It also provides a higher degree of relevance with reference to its surrounding and properly differentiates from the nonadjacent objects. Furthermore, it is clear from Table 1 that for testing purposes, the state-of-the-art systems have considered a small testing dataset. However, for result evaluation of a proposed system, we have requested three surgeons to choose the ground truth from our dataset. A ground truth is selected in terms of video, a relevant approach, and a mechanism can also be found in [9]. Sample frames from meniscal surgery are shown in Figures 6–8.

3.2. Tool. To precede the evaluation phase, an input video is provided to our proposed system. Its spatiotemporal features are computed against 1–N targeted surgical videos, and the most relevant videos are displayed in the right side of the interface as shown in Figure 9.

3.3. Evaluation Measurement. For the purposes of a system evaluation, there are many parameters such as precision,

Figure 6: Frame extracted from reconstruction surgery of meniscal (1).

Figure 7: Frame extracted from reconstruction surgery of meniscal (2).

Figure 8: Frame extracted from reconstruction surgery of meniscal (3).

recall, and ROC area. But in the information retrieval system, precision metric is used because it provides the proportion of retrieved videos that are actually relevant. In order to measure the precision of our system, we have split the video into small segments and these segments are further compared with the ground truth. Precision can be computed using the following metric [26, 27].

$$\text{Precision} = \frac{\text{TP}}{\text{TP}} + \text{FP}. \tag{6}$$

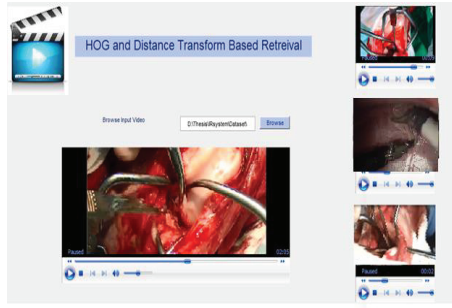


Figure 9: Sample screenshot of relevant videos using HOG and CFD.

Herein, a video is said to be the true positive (TP) if it is selected by both our system and the surgeon, or said to be false positive (FP) if it is selected by our system but not by the surgeon. Using this scheme, we measure the average precision of our system using our four proposed methods as shown in Table 2. It has already been mentioned that video retrieval system specifically for meniscus surgery was not yet found in the literature. Therefore, we have considered a number of general parameters for video retrieval system for result comparison. A brief comparison of medical video retrieval system and our proposed system is shown in Table 2.

We can deduce from Table 2 that PCA-SIFT from our proposed method performs well. Observable key factors of this dominant performance are that the incline areas nearby SIFT points are extremely organized to be presented through PCA. Eigenvalues against the patch decline are considerably faster than Edigenvalues for arbitrarily designated incline patches. For the reason that patches adjoining points entirely share definite characteristics, reducing since SIFT point recognition stage; they are totally adjusted, interchanged to align central inclines with vertical, and are scaled correctly. That streamlines the trade PCA necessity do and enabling the PCA-SIFT to exactly characterize patch by the small number of dimensions along with Lucas-Kanade optical flow method.

4. Conclusion

Our proposed video retrieval system for meniscal surgery modeling is based on four methods. These methods include key point-based retrieval, statistical retrieval, PCA-SIFT-based retrieval, and PCA-GMM-based retrieval. They further exploit spatiotemporal features for the improvement of health services to human beings.

In a key point-based retrieval method, key points of video are extracted using SIFT, HOG, CFD, and DT. Secondly, using the statistical retrieval method, spatial and temporal attributes are statistically modeled using GMM and Kullback-Leibler divergence. Moreover, by using PCA-SIFT-based retrieval and PCA-GMM-based retrieval methods, the dimension reduction technique is used with SOFT and GMM, respectively. To track spatiodevelopment over temporal interims, Lucas-Kanade sparse-optical flow method was used, which is efficient and robust to noise.

Table 2: Comparative analysis of proposed system with already existing systems.

	Dataset	Clip size	Average retrieval time	Average number of match frames/ videos	Recall	Average precision	ROC	Description
1	Medical videos & classified human action	Avg 20 sec	7 min 3 sec	0	0%	0.43	0	*Time* required to compute feature vector for 9 videos 0.43 in case of ERM and CD dataset
2	Medical videos	1.97 min and above	16.4 (retrieval time for all attributes)	111	0%	0.983	0	*Precision* is based on 8 videos using the neural network in terms of relevant key frame retrieval and a total number of key frames for that video
3	Laparoscopic videos		0	0	80%	0	0	N/A
4	Endoscopic videos	60 sec to 250 sec	0	0	88%	0	0	N/A
5	Endoscopy	0	0	0	0%	0.78	0	N/A
6	Cataract	0	0	0	0%		0.7	With motion analysis and 0.9 with the presence of surgical process
7	Hysteroscopy	2 to 3 min	0	0	0%	0.92	0	
8	Eye surgery	0	0	0	0%	0.72	0	In the case of CSD dataset
9	Epiretinal surgery	0	0	0	0%	0.62	0	0.69 in case of motion histogram
Proposed	Meniscal surgery	60 sec	60 sec	10	78%	0.57 0.68 0.64 0.78	0	0.78 for PCA-SIFT

For experimental work, we have used 20 videos of meniscus surgery for human health services, and our evaluation shows that results gained using PCA-SIFT are more significant. The results established using PCA-SIFT approach enhanced the obvious results with an average precision of 0.78.

Our proposed system will help out the medication of diverse patients, and surgeons can especially get guidance by watching the videos of surgeries performed by experts. Various datasets of real meniscal surgery in our system will be explored in the future for testing purposes and evaluation.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by the MIST (Ministry of Science and ICT), Korea, under the National Program for Excellence in SW supervised by the IITP (Institute for Information & Communications Technology Promotion) (2015-0-00938).

References

[1] J. B. Thorlund, K. B. Hare, and L. S. Lohmander, "Large increase in arthroscopic meniscus surgery in the middle-aged and older population in Denmark from 2000 to 2011," *Acta Orthopaedica*, vol. 85, no. 3, pp. 287–292, 2014.

[2] K. Charriere, G. Quellec, M. Lamard, G. Coatrieux, B. Cochener, and G. Cazuguel, "Automated surgical step recognition in normalized cataract surgery videos," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 4647–4650, Chicago, IL, USA, 2014.

[3] P. Over, J. Fiscus, G. Sanders et al., "TRECVID 2014 – an overview of the goals, tasks, data, evaluation mechanisms and metrics," in *Proceedings of TRECVid*, p. 52, Orlando, FL, USA, 2014, NIST.

[4] S. T. Ramya, P. Rangarajan, and V. Gomathi, "XML based approach for object-oriented medical video retrieval using neural networks," *Journal of Medical Imaging and Health Informatics*, vol. 6, no. 3, pp. 794–801, 2016.

[5] K. Schoeffmann, C. Beecks, M. Lux, M. S. Uysal, and T. Seidl, "Content-based retrieval in videos from laparoscopic surgery," in *Medical Imaging 2016: Image-Guided Procedures, Robotic Interventions, and Modeling*, vol. 9786, pp. 1–10, San Diego, CA, USA, 2016.

[6] C. Beecks, K. Schoeffmann, M. Lux, M. S. Uysal, and T. Seidl, "Endoscopic video retrieval: a signature-based approach for linking endoscopic images with video segments," in *2015 IEEE International Symposium on Multimedia (ISM)*, pp. 33–38, Miami, FL, USA, 2015.

[7] J. R. Carlos, M. Lux, X. Giro-i-Nieto, P. Munoz, and N. Anagnostopoulos, "Visual information retrieval in endoscopic video archives," in *2015 13th International Workshop on Content-Based Multimedia Indexing (CBMI)*, pp. 1–6, Prague, Czech Republic, 2015, IEEE.

[8] K. Charriere, G. Quelled, M. Lamard et al., "Real-time multilevel sequencing of cataract surgery videos," in *2016 14th International Workshop on Content-Based Multimedia Indexing (CBMI)*, pp. 1–6, Bucharest, Romania, 2016, IEEE.

[9] K. Muhammad, M. Sajjad, M. Y. Lee, and S. W. Baik, "Efficient visual attention driven framework for key frames extraction from hysteroscopy videos," *Biomedical Signal Processing and Control*, vol. 33, pp. 161–168, 2017.

[10] G. Quellec, M. Lamard, Z. Droueche, B. Cochener, C. Roux, and G. Cazuguel, "A polynomial model of surgical gestures for real-time retrieval of surgery videos," in *Medical Content-Based Retrieval for Clinical Decision Support*, vol. 7723 of Lecture Notes in Computer Science, , pp. 10–20, Springer, 2013.

[11] C. A. Kelsey, P. H. Heintz, D. J. Sandoval, G. D. Chambers, N. L. Adolphi, and K. S. Paffett, "Radiation interactions with tissue," in *Radiation Biology of Medical Imaging*, pp. 61–79, John Wiley & Sons, Inc., 2014.

[12] D. Markonis, M. Holzer, F. Baroz et al., "User-oriented evaluation of a medical image retrieval system for radiologists," *International Journal of Medical Informatics*, vol. 84, no. 10, pp. 774–783, 2015.

[13] S. W. Zhang, Y. J. Shang, and L. Wang, "Plant disease recognition based on plant leaf image," *The Journal of Animal & Plant Sciences*, vol. 25, no. 3, Supplement 1, pp. 42–45, 2015.

[14] J. Choi, Z. Wang, S. C. Lee, and W. J. Jeon, "A spatio-temporal pyramid matching for video retrieval," *Computer Vision and Image Understanding*, vol. 117, no. 6, pp. 660–669, 2013.

[15] H. Bai, Y. Dong, S. Cen et al., *Orange Labs Beijing(FTRDBJ) at TRECVid 2013: Instant Search*, International Journal of Multimedia Information Retrieval, 2013.

[16] S. Bharathi, D. Shenoy, R. Venugopal, and M. Patnaik, "Ensemble PHOG and SIFT features extraction techniques to classify high resolution satellite images," *Data Mining and Knowledge Engineering*, vol. 6, no. 5, pp. 199–206, 2014.

[17] A. Bruhn, J. Weickert, and C. Schnörr, "Lucas/Kanade meets Horn/Schunck: combining local and global optic flow methods," *International Journal of Computer Vision*, vol. 61, no. 3, pp. 1–21, 2005.

[18] Z. Droueche, M. Lamard, G. Cazuguel, G. Quellec, C. Roux, and B. Cochener, "Motion-based video retrieval with application to computer-assisted retinal surgery," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 4962–4965, San Diego, CA, USA, 2012.

[19] Z. Droueche, G. Quellec, M. Lamard, G. Cazuguel, B. Cochener, and C. Roux, "Computer-aided retinal surgery using data from the video compressed stream," *International Journal of Image and Video Processing: Theory and Application*, vol. 1, no. 1, 2014.

[20] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[21] B. V. Patel and B. B. Meshram, "Content based video retrieval systems," *International Journal of UbiComp*, vol. 3, no. 2, pp. 13–30, 2012.

[22] B. Werner, S. Harald, and M. Roland, *Semantic Indexing and Instance Search; Joanne Research at TRECVid*, TRECVID Workshop, 2013.

[23] T. Tuytelaars and K. Mikolajczyk, "Local invariant feature detectors: a survey," *Foundations and Trends® in Computer Graphics and Vision*, vol. 3, no. 3, pp. 177–280, 2008.

[24] MedlinePlusU.S. National Library of Medicine, Bethesda, MD, USA, April 2017 https://medlineplus.gov/surgeryvideos.html.

[25] American Academy of Orthopaedic SurgeonsApril 2017, http://www.aaos.org/videogallery/.

[26] J. Davis and M. Goodrich, "The relationship between precision-recall and ROC curves," in *Proceedings of the 23rd international conference on Machine learning - ICML '06*, pp. 233–240, Pittsburgh, PA, USA, 2006.

[27] M. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," 2014, March 2016, https://medtube.net/orthopedics.

Hindawi
Journal of Sensors
Volume 2018, Article ID 3265281, 13 pages
https://doi.org/10.1155/2018/3265281

Research Article

Application of a Mobile Chronic Disease Health-Care System for Hypertension Based on Big Data Platforms

Dingkun Li,[1] **Hyun Woo Park**,[1] **Erdenebileg Batbaatar**,[1] **Lkhagvadorj Munkhdalai**,[1] **Ibrahim Musa**,[1] **Meijing Li**,[2] **and Keun Ho Ryu**[1]

[1]*Database/Bioinformatics Lab, School of Electrical & Computer Engineering, Chungbuk National University, Cheongju 28644, Republic of Korea*
[2]*College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China*

Correspondence should be addressed to Keun Ho Ryu; khryu@chungbuk.ac.kr

Received 15 February 2018; Accepted 26 March 2018; Published 29 May 2018

Academic Editor: Ka L. Man

Hadoop is a globally famous framework for big data processing. Data mining (DM) is the key technique for the discovery of the useful information from massive datasets. In our work, we take advantage of both platforms to design a real-time and intelligent mobile health-care system for chronic disease detection based on IoT device data, government-provided public data and user input data. The purpose of our work is the provision of a practical assistant system for self-based patient health care, as well as the design of a complementary system for patient disease diagnosis. This system was only applied to hypertensive disease during the first research stage. Nevertheless, a detailed design, an implementation, a clear overview of the whole system, and a significant guide for further work are provided; the entire step-by-step procedure is depicted. The experiment results show a relatively high accuracy.

1. Introduction

Hypertension is a condition in which a person's blood pressure is above normal or optimal limit of 120 mmHg for systolic pressure and 80 mmHg for diastolic pressure. Increased blood pressure in the long term can lead to conditions that could threaten the health of the sufferer. Several conditions can cause disturbances in hypertensive cardiovascular organs such as stroke and heart failure, so sometimes mentioned that hypertension is a silent killer, because sufferers sometimes do not realize that he was exposed to conditions of hypertension [1]. The classification of blood pressure in adult divided into 4 classes which have been shown in Table 1.

Nevertheless, in our work, we treat prehypertension, HP stage 1, and HP stage 2 as hypertension with no difference.

Hadoop, which is based on MapReduce, has been one of the most important and popular techniques in the field of big data analysis during the last few years. Undoubtedly, it is the key technique for massive data analysis. Alternatively, Spark is a promising distributed framework that runs memory data on clusters at a speed that is considerably faster than that of Hadoop. Data mining (DM) is the key technique for the discovery of the useful information from well-processed data at the intersection of areas such as machine learning, statistics, and database systems. In the present work, the aim is the exploitation of the use of Hadoop, Spark, and DM techniques to provide a more powerful way of handling big data at high extents of speed, safety, and accuracy.

In recent years, the Hadoop framework has been widely used for the delivery of health care as a service [2]; moreover, a wide variety of organizations and researchers have used Hadoop for health-care services and clinical-research projects [3]. Taylor provided a detailed introduction on the use of Hadoop in bioinformatics [4], while Schatz developed an operations support system (OSS) package named CloudBurst

Table 1: Classification of Blood Pressure.

Classification	Systolic (mmHg)	Diastolic (mmHg)
Normal	<120	And <80
Prehypertension	120 ~ 139	Or 80 ~ 89
HP Stage 1	140 ~ 149	Or 90 ~ 99
HP Stage 2	≥150	Or ≥100

that provides an algorithmic parallelization model for which Hadoop MapReduce is used [5]. Indeed, the Hadoop framework has been employed in numerous important works to provide major contributions to the health-care field. The other big data processing framework, Spark, leverages a synergistic combination of the smartphone and the smartwatch in the monitoring of multidimensional symptoms such as facial tremors, dysfunctional speech, limb dyskinesia, and gait abnormalities [6].

Over many years, a large amount of health-care research work has been completed using DM techniques. In [7, 8], the authors used classification and regression techniques to predict conditions like cardiovascular disease and heart disease. In [9, 10], integrated DM techniques are provided for the detection of chronic and physical diseases. Further, a number of other research works, like [11, 12], used the advantages of DM to develop new methodologies and frameworks for health-care purposes.

The major goal of health-informatics research is the improvement of the quality and the cost of care that are provided to users, or the health-care output [13]. The purpose of the present work is the exploitation of Hadoop, Spark, and DM techniques for the design of a comprehensive, real-time, and intelligent mobile health-care system for chronic disease detection and prediction. The system is designed to provide an assistant system for self-based user health care, as well as a complementary system for the daily diagnostic work of doctors.

A series of challenges arise in the development of a big data-based health-care system. Firstly, it is extremely difficult to obtain high-quality and relevant medical data. One reason for this is that hospitals or the patients themselves are not willing to offer personal data for public research due to privacy policies. Another reason is the need to engage with a variety of data sources, such as the collection of data from hospitals, health-care centers, governments, laboratories, and the patients' families, which can cause serious missing-data problems. For instance, only the hospital-treatment data of patient A are available while the lifestyle (smoking, drinking, etc.) data are missing, and only the lifestyle data of patient B are available while the patient's treatment data are missing. The work of [14] confirms this varied-source characterization of health-care data collection and the complexity of different data forms. Secondly, data analysis is a challenging work. Even though a great quantity of research work has been completed to process and analyse data, a high-quality framework with highly precise predictive and analytic results is still mostly elusive [1]. Thirdly, the difficulty regarding the creation of a tool that can break the borders between the patient, health-care providers, and public health-care organizations to connect these parties in a practically meaningful manner is another obstacle [15].

The contributions of the present work are as follows: (1) exploration of the possibility of the utilization of Hadoop, Spark, and DM techniques in the work regarding health-care big data. (2) Depiction of a detailed step-by-step design of the health-care system for disease detection and prediction. (3) Provision of an overview for the next research stage and a guide for other similar systems. (4) Minimization of the monetary cost through the use of the Google Cloud services FCM and GCSql, which also guarantee real-time data transactions. A preliminary version of the present work has been reported in [15].

This paper is organized as follows: a description of the related work is provided in Section 2; an overview of the proposed system is introduced in Section 3; the design details are described in Section 4; the experiment results are described in Section 5; and Section 6 concludes this work and introduces future work.

2. Selection Techniques and Algorithms

This section briefly describes the related platforms, algorithms, and some of the key techniques that were used in the undertaking of the present work.

2.1. Hadoop, Spark, and Data Mining. Hadoop consists of the HDFS (Hadoop Distributed File System), HBase, and Hadoop MapReduce, making it very suitable for big data analyses [16]. As a 100% open-source framework, it has been widely used in almost every field for big data processing. In the last few years, Apache Spark [17] received great attention in the big data and data science fields, mainly because of its easier, friendlier application program interface (API) and an enhanced memory management compared with MapReduce; therefore, developers could concentrate on the data-computation logical operations rather than the background details of the computational execution.

It is difficult to find a coincident DM definition, but one of the widely accepted definitions states that DM is the process of discovering interesting patterns and knowledge from large amounts of data [18]. Its other close concept is called knowledge discovery in databases (KDD); DM is the analytical step of KDD. In this paper, a commonly used classification algorithm called C4.5 will be used for the disease-rule generation since it is simple, stable, and produces results of a relatively high accuracy.

2.2. C4.5. C4.5 is an algorithm that was developed by Ross Quinlan and is used to generate decision trees [18]. C4.5 is an extension of Quinlan's earlier ID3 algorithm. The decision trees that are generated by C4.5 can be used for the purpose of classification, and for this reason, C4.5 is often referred to as a statistical classifier.

In general, the steps of the C4.5 algorithm for the building of decision trees are as follows: choose the attribute for the root node; create the branch for each value of that attribute; split the case according to the branches; and repeat

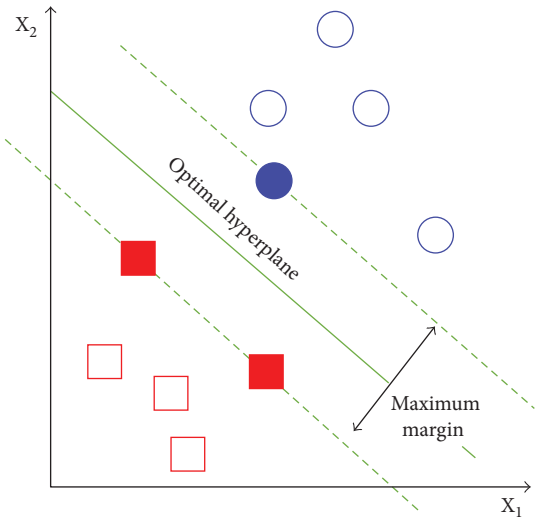


Figure 1: Maximum margin, the vectors on the dashed line are the support vectors.

the process for each branch until all of the branch cases are of the same class [18].

2.3. Support Vector Machine. Support vector machine (SVM) [19] has been used to select features and generate the classifier. For feature selection, this method is a backward sequential selection approach. One starts with all the features and removes one feature at a time until only r features are left. The operation of the SVM algorithm is based on finding the hyperplane that gives the largest minimum distance to the training examples. The basic concept is described using Figure 1.

The strategy ranks the features according to their influence on the decision hyperplane.

The optimal hyperplane is used to classify the data into different classes in two or more dimensionalities.

2.4. Hybrid Feature Selection Mechanism. Feature selection aims at finding the most relevant features of a problem domain. Primarily, there are two kinds of feature selection methods, filters and wrappers. The filters work fast but its result is not always satisfactory. While the wrappers guarantee good results, they are very slow when applied to wide feature sets which contain hundreds or even thousands of features. According to work [20], a hybrid feature selection mechanism takes advantage of both filter and wrapper feature selection methods is used to improve the computation speed and accuracy.

Inspired by [20], we developed our feature selection mechanism. The architecture is show in Figure 2.

3. Main Framework

An overview of the whole system is given in this section, and this is followed by a description of the implementation details in Section 4. The proposed system comprises four modules. The overview of the architecture of the entire system is depicted in Figure 3 [15]. The four modules in the figure

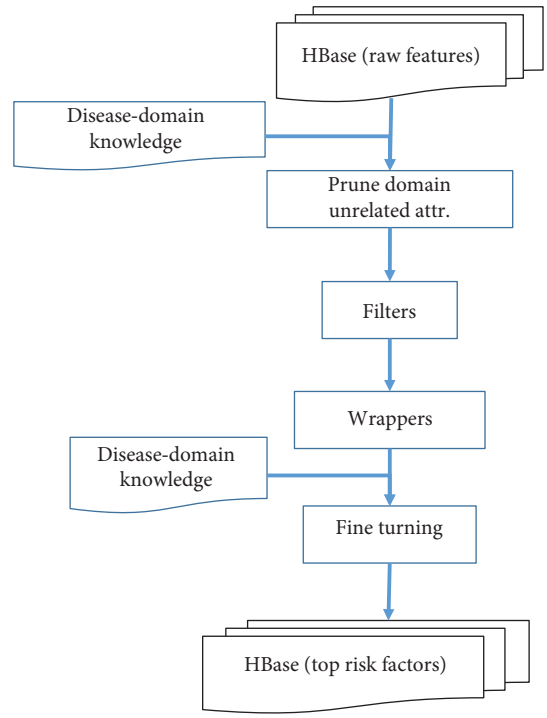


Figure 2: Risk factors (features) selection mechanism. We combined the expert domain knowledge for the purpose of pruning the unrelated attributes. Filter methods followed by wrapper methods are used select the features for disease rule generation. The evaluated result is stored in HBase.

are named as follows: (1) data collection module, (2) data storage module, (3) third-party-server (TPS) module, and (4) Cloud service module.

Module 1a is used to collect the streaming data and structured data by IoT devices such as Fitbit Charge 2, mobile phone sensors. 1b is used to import structured, semistructured, and unstructured data from various data sources like hospitals, governments, families, user inputs, and so on. Besides, we have developed a mobile app to collect user input data such as lifestyle and food intake data.

Module 2 is used to store the data in HBase collected by module 1. The data collected by the system is of three types: the structured, the semi-structured, and the unstructured data. Firstly, all these three kinds of data will be stored in HBase as it is quite suitable for mass data preprocessing and storage. Then this data should be converted into structured data for further processing.

Module 3 is used for the processing and analysis of the data based on the Hadoop/Spark cluster which is the key module of the whole system; all the data processing and analysis work will be done by this module. It is used for data statistical analysis, patient emergency detection, and disease prediction and detection. It also responses for message like data analysis results generation. These result will be sent to module 4.

Module 4 is used for the message dissemination. This model is implemented by using Google Cloud SQL (GCSql) and Google Firebase Cloud Messaging (GFCM) services. When receiving the requests from the TPS, Cloud model

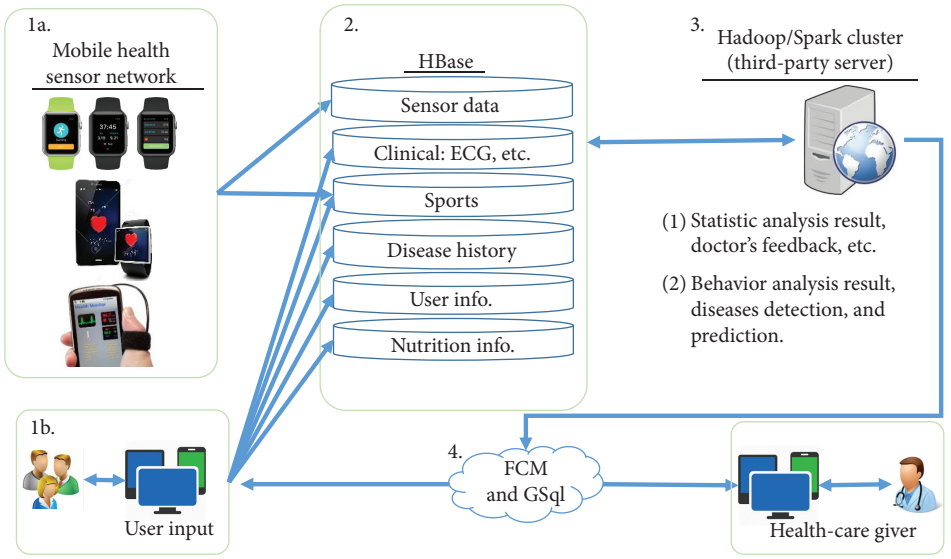


Figure 3: Systemic architecture: the four modules are marked 1 (1a and 1b), 2, 3, and 4.

responses immediately according to these requests, stores data, or sends data to the devices registered to it.

Further details have been given in a previous work of the authors of the present study [15].

4. System Implementation Details

In this section, descriptions of the systemic data flow, the data storage and processing, and the disease detection and prediction based on large medical datasets are provided.

4.1. Data Collection, Preprocessing, and Storage. To obtain high-quality structured datasets, database processing, natural language processing (NLP), and image-processing techniques are combined with the DM data-preprocessing techniques that are used by the TPS to process the different kinds of data (structured, semistructured, and unstructured), and the data are then transformed into a structured data record. The result is then stored in the HBase.

(1) For the structured dataset (mostly imported from the other public Web services) that includes patient information, prescriptions, and disease histories, it is relatively easier to import the data from the rational DB to the HBase using Sqoop [21].

(2) For the semistructured dataset, which includes HTML, XML, and Json documents, the TPS will design row keys like the d001 for the HBase table, including its document-information column value together with its family map that is called "column family" (it comprises the document timestamps of the HBase), as shown in Figure 4. The semistructured data will be converted to the structured data in the HBase, as shown in Figure 5.

(3) For the unstructured data, like clinic notes and the stream data from mobile sensors, they will be managed by the system in a particular way. The clinic notes contain a lot of the textual information, [22] providing an efficient way to convert this data into structured data through the use of NLP techniques, text-mining algorithms, and the MapReduce framework. The same strategy is used in the proposed system to deal with this problem, and the procedure is shown in Figure 6.

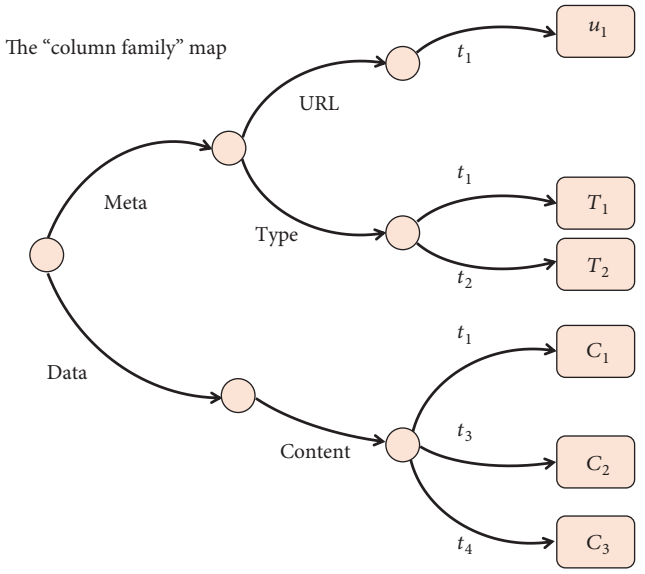


Figure 4: Column family map. The meta information and data content can be depicted by this map with their timestamp information. Where u_i stands for document URL, T_i stands for document type, and C_i stands for document content.

The stream data are handled using Apache Spark [23] techniques; the basic procedure is shown in Figure 7. Finally, the output will be stored in the HBase. After the preprocessing step, all kinds of data will be converted into structured data and stored in distributed HBase regional servers for further processing.

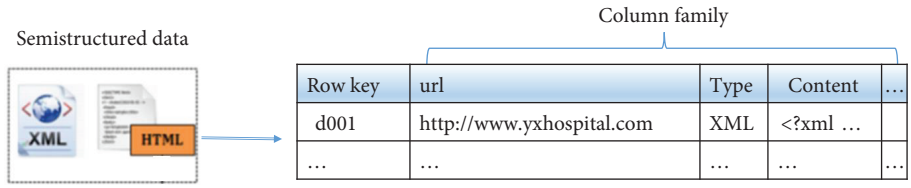


Figure 5: HBase example for semistructured data storage. Semistructured data (such as XML and HTML files) will be converted to structured data which has the similar structure as relational data base table.

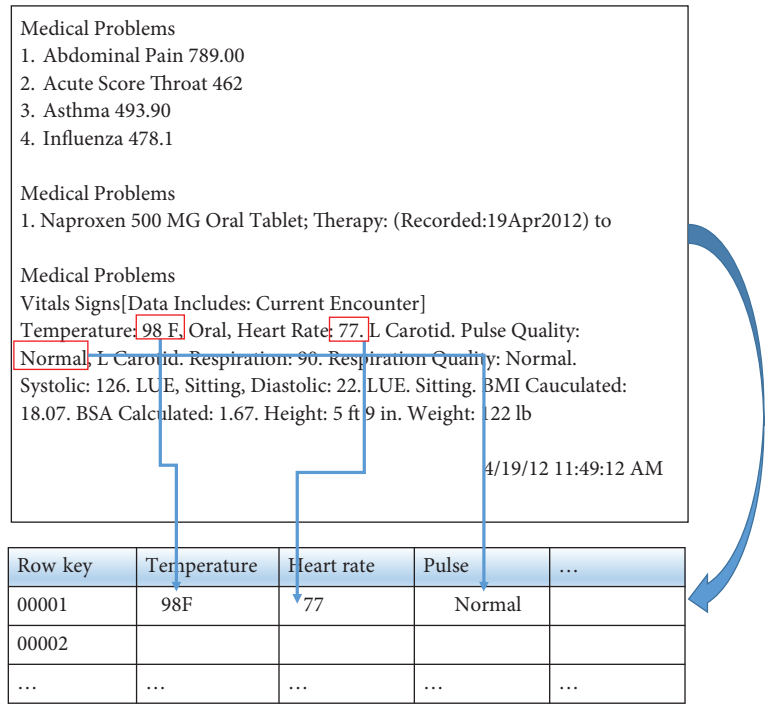


Figure 6: HBase example for unstructured data storage. Important information will be extracted from the raw data and stored in table format which is appropriate for further analysis.

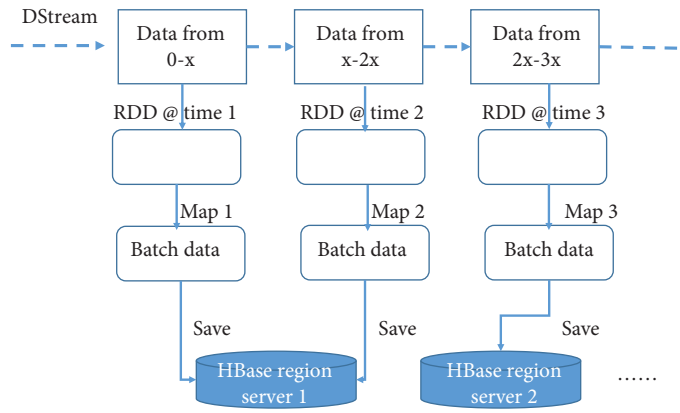


Figure 7: Streaming-data processing by Spark. The Spark engineer divides the stream data into sequential blocks and these blocks are stored in different nodes of the cluster.

4.2. Disease Data Statistical Analysis. Among the whole existing dataset, some of the patient data are treated as the training set for the disease-rule generation (some of the datasets are not disease related). The first step here is the counting of the diseases of the patients with their personal information, like their gender, age, nationality, and occupation. Since the data were stored in the HBase, the MapReduce framework was used to count the diseases. The training data

Condition attribute value | Decision attribute value

Row key	Temperature	Heart rate	Pulse	…	Disease
00001	98F	77L	Normal		Hypertension
00002	98.5F	80L	Normal		Diabetes I
00003	97F	70L	Fast		Diabetes II
00001	99.5F	79L	Normal		Bradycardia
…	…	…	…	…	…
00008					Hypertension
…	…	…	…	…	…

HBase region server 1
HBase region server 2
……

Figure 8: Training data example. Data comes from different sources and stored in different cluster nodes.

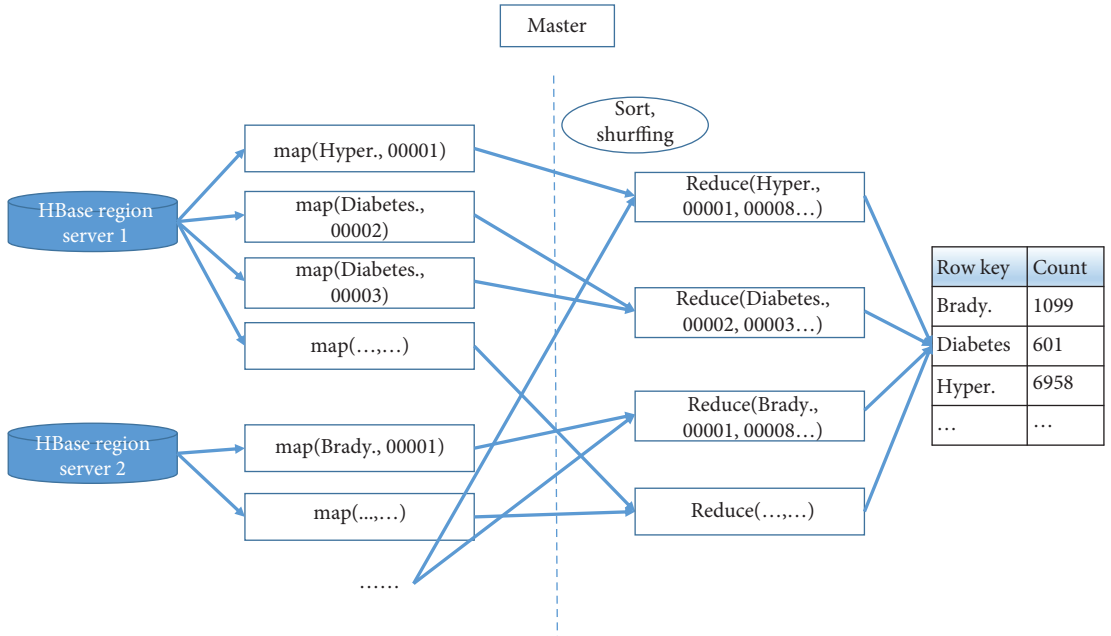


Figure 9: MapReduce procedure for disease count. Disease name is key and row ids are value for Map function. Disease name is key and all the row ids for the same disease are value for Reduce function. The count of row ids is the count for specified disease.

are stored in separate regional servers, as shown in Figure 8. The disease-count MapReduce procedure running in TPS is depicted in Figure 9.

The disease-count algorithm running in distributed environment is shown in Algorithm 1. It consists of two main procedures which are Map and Reduce. Map function is used to assign constant value 1 in terms of each row for distinct disease and patient. Reduce function is used to add all 1s together for the same disease, the sum of 1s is the count of the specified disease.

Based on the output, it is straightforward to obtain the patient list for a specified disease, as well as all of the personal information according to the patient ID.

4.3. Risk Factor Selection. The risk factor (RF) selection procedure is a process for feature selection. Hybrid feature selection has been applied to the raw dataset. First, we apply t-test

Input:
HBase table
Output:
Diseases count and related info
1. **class** Mapper
2. **method** map (HBase table)
3. **for** each instance row in table
4. **write** ((disease$_i$, patientID), 1)
5.
6. **class** Reducer
7. **method** reduce ((disease$_i$, patientID), ones[1,1,1,…n])
8. sum=0
9. **for each** one in ones do
10. sum+=1
11. **return** ((disease$_i$, patientID),sum)

Algorithm 1: Disease count running on HBase.

Input:
Hypertension disease data set
Output:
Selected features and SVM classifier
1. **load** data set
2. **Sample** the data randomly into training (67%) and testing(33%) data set
3. **set** target variable
4. generate the classifier based on the training data set
5. train the classifier using linear kernel function (similarity function)
6. predict the testing data set using the trained classifier
7. evaluate the classifier
8. Recursively select the features correspond to their weights

Algorithm 2: Linear SVM pseudocode.

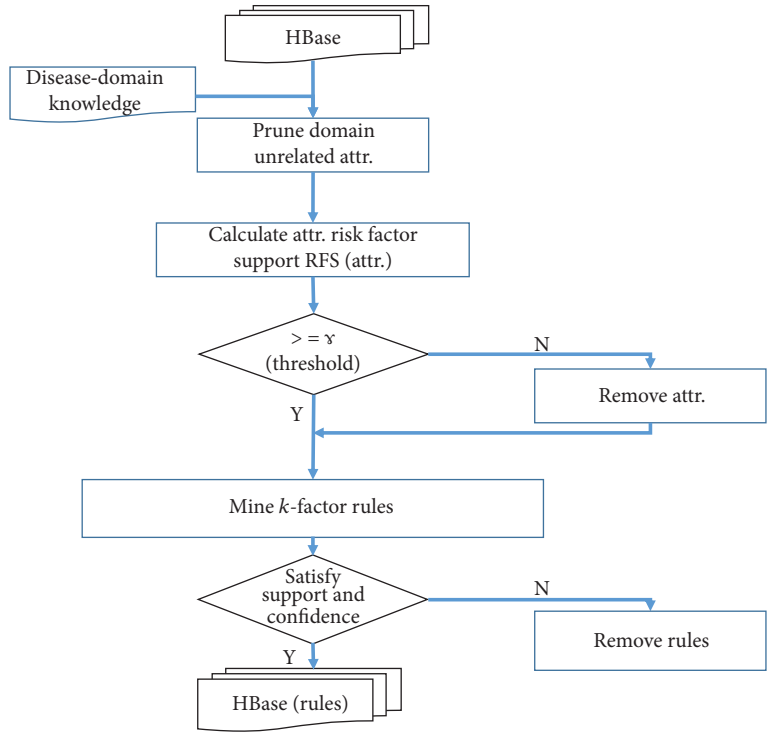


Figure 10: Disease-rule generation procedure. Domain expert's knowledge is used to prune the unrelated attributes for the first round. Then RFS is calculated for each remaining attributes to compare with predefined threshold γ, if larger than γ, it is treated as risk factor for further rule generation step.

combined with chi-squared test as filters to prune unrelated features according to research [11], chi-square and t-test are fast and can achieve relative high accuracy. The chi-squared test formula is,

$$\chi^2 = \sum_{i=1}^{k} \frac{(x_{i-} m_i)2}{m_i}, \tag{1}$$

where the expected numbers m_i were large enough known numbers in all cells assuming every x_i may be taken as normally distributed, and reached the result that, in the limit as n becoming large, χ^2 followed the chi-squared distribution with (k-1) degrees of freedom.

The t-test can be used, for example, to determine if two sets of data are significantly different from each other. The t-test

$$t = \frac{Z}{s} = \frac{(\bar{x} - \mu)/(\sigma/\sqrt{n})}{s}, \tag{2}$$

where $\bar{x}$ is the sample mean from a sample $X_1, X_2, \ldots, X_n$, of size n, s is the ratio of sample standard deviation over population standard deviation, σ is the population standard deviation of the data, and μ is the population mean.

After the first step, we propose a wrapper method based on linear support vector machines (SVMs). Firstly, we train the linear SVM on a subset of training data and retain only

Input:
Data partition D: a training set and associated class label C
Attribute list L (selected disease risk factors in previous step)
Output:
Decision tree with its root N
Method:
1. Create a node N,
2. **if** samples has the same class, C then,
3. **return** N as leaf node with class C label
4. **if** list of attributes is empty then
5. **return** N as leaf node with class label that is the most class in training set.
6. Choose test factor, that has the most GainRatio using attribute_selection_method
7. give node N with test-attribute label
8. **for each** attribute a_i in L
9. add branch in node N to test-attribute=a_i
10. make partition for sample s_i from training set where test-attribute=a_i
11. **if** s_i is empty then
12. attach leaf node with the most class in training set
13. **else** attach node that generated by Gnerate_decision _tree (s_i, L, test-attribute)
14. **return** N

Algorithm 3: C4.5 for important disease rule generation.

Table 2: Disease-count results.

Disease	Hyperten.	Dyslip.	Ostarthritis.	Diabetes	Asthm.
Count	9385	3213	5223	3500	923

Table 3: Statistical analysis.

	Weighted *N* (%)	Women (%)	Men (%)
Age			
20–44	20,131 (40.6)	11,416 (40.5)	8715 (41.5)
45–64	17,623 (35.6)	9946 (35.3)	7677 (36.6)
65+	11,775 (23.8)	6822 (24.2)	4593 (21.9)
Marital status			
Married	42,478 (86.6)	24,747 (88.5)	17,731(84.1)
Never married	6569 (13.4)	3208 (11.5)	3361 (15.9)
Education			
University	12,990 (28.9)	6612 (25.5)	6378 (33.6)
High school	14,876 (33.1)	8129 (31.3)	6747 (35.6)
Middle school	4950 (11.0)	2659 (10.2)	2291 (12.1)
Elementary school	12,125 (27.0)	8568 (33.0)	3557 (18.8)
Occupation			
Office worker	8993 (20.1)	4104 (15.8)	4889 (25.9)
Manual worker	17,181 (38.4)	8099 (31.3)	9082 (48.2)
Unemployed	18,602 (41.5)	13,710 (52.9)	4892 (25.9)
Income Level			
1st quartile	12,077 (24.8)	6862 (24.8)	5215 (24.9)
2nd quartile	12,183 (25.1)	6940 (25.1)	5243 (25.0)
3rd quartile	12,193 (25.1)	6930 (25.1)	5263 (25.1)
4th quartile	12,167 (25.0)	6912 (25.0)	5255 (25.1)
Hypertension			
No	8619 (47.9)	4954 (47.4)	3665 (48.5)
Yes	9383 (52.1)	5488 (52.6)	3895 (51.5)

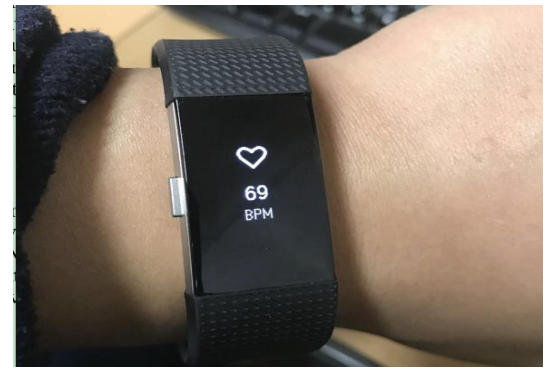


Figure 11: Screenshot of the IoT device we used in our system.

those features that correspond to highly weighted components (in absolute value sense) of the normal to the resulting hyperplane that separates positive and negative examples for the class. Secondly, recursively eliminate features whose weight value is close to zero. Finally, the features remained are selected as risk factor candidates. We created a representation result in experimental session, Chapter 5. The pseudocode is given in Algorithm 2 below. First, we divide the original dataset into training and testing dataset; the SVM classifier is generated based on this training dataset. After evaluating the classifier, the features will be recursively selected according to the weight until the stop criterion has been met.

The result will be described in experimental session.

4.4. Disease-Rule Generation. The format of disease rules is like that of the IF THEN rule; for example, IF (edu = elementary, B1 < =0.86 mg/day, married), THEN (hypertension = yes). The purpose here is the mining of all of the disease-related rules from the training dataset for a further data analysis including disease prediction and disease

region town_t apt_t sex age incm ho_incm edu occp kstrata cfam genertn allownc house live_t ainc_unit1
ainc marri_1 marri_2 fam_rela tins npins M_1_yr mt_nontrt BH9_11 BH9_13 BH1_1 BH1_2 BH1_3
BH1_8 BH1_6 BH2_61 BH2_62 BH2_63 BH2_66 BH2_67 BH2_64 LQ4_01 LQ4_02 LQ4_03
LQ4_04 LQ4_05 LQ4_06 LQ4_07 LQ4_08 LQ4_09 LQ4_10 LQ4_11 LQ4_12 LQ4_13 LQ4_14
LQ4_15 LQ4_16 LQ4_21 LQ4_22 LQ4_23 LQ4_17 LQ4_18 LQ4_19 LQ4_20 LQ1_sb LQ2_ab
LQ2_mn LQ_1EQL LQ_2EQL LQ_3EQL LQ_4EQL LQ_5EQL EQ5D graduat EC1_1 EC1_2 EC_occp EC_stt_1 EC_stt_2
EC_wh EC_wht_0 EC_wht_23 EC_wht_5 EC_lgw_2 EC_lgw_4 EC_lgw_5 EC_pedu_1 EC_pedu_2 BO1 BO1_3 BO2_1
BO3_01 BO3_02 BO3_03 BO3_14 BO3_05 BO3_04 BO3_12 BO3_07 BO3_09 BO3_10 BD2 BD1_11
BD2_1 BD2_31 BD2_32 BD7_4 BD7_6 BD7_5 dr_month BA2_12 BA2_13 BA1_3 BA1_5
BA2_2_1 BA2_2_2 BA2_2_5 BA2_2_6 BA2_22 sc_seatblt2 BP8 BP1 BP7 mh_stress BS1_1 BS3_1 BS3_3 BS6_2 BS6_2_1
BS6_2_2 BS6_3 BS6_4 BS6_4_1 BS6_4_2 BS5_2 BS5_21 BS5_22 BS5_24 BS5_26 BS5_28 BS5_32 BS5_25
BS5_27 BS5_29 BS5_30 BS8_2 BS9_1 BS9_2 BS13 BS12_1 BS12_2 BE3_71 BE3_73 BE3_74 BE3_81
BE3_82 BE3_83 BE3_84 BE3_91 BE3_92 BE3_93 BE3_94 BE3_75 BE3_76 BE3_77 BE3_78
BE3_85 BE3_86 BE3_87 BE3_88 BE8_1 BE8_2 BE3_31 BE3_32 BE3_33 BE5_1 BE5_2 pa_aerobic LW_ms
LW_mp_a LW_ms_a LW_pr LW_pr_1 LW_mt LW_mt_a1 LW_mt_a2 LW_br LW_br_ch LW_br_dur LW_br_yy
LW_br_mm HE_HP O_DMFTP O_DMFIP OR1 O_ortho BM1_0 BM1_1 BM1_2 BM1_3 BM1_4 BM1_5
BM1_6 BM1_7 BM1_8 BM2_3 BM2_2 BM2_4 BM2_5 BM13 BM7 O_chew_d BM8 OR1_2 MO4_00
BM12 BM12_1 T_Q_SNSTDG T_Q_SNSTPT T_Q_OMPT T_Q_HR T_Q_HR1 T_Q_VN T_Q_VN1 T_NQ_PH
T_NQ_PH_T T_NQ_OCP T_NQ_OCP_T T_NQ_OCP_P T_NQ_LS T_NQ_LS_T T_NQ_FIR T_NQ_FIR_P GS_use GS_mea_r_1
TH_ult TH_ult_1 TH_ult_2 TH_deli_1 L_BR L_DN L_BR_FQ L_DN_FQ L_OUT_FQ L_BR_TO L_BR_WHO L_LN_TO
L_LN_WHO L_DN_TO L_DN_WHO LK_LB_CO LK_LB_US LK_LB_IT LK_LB_EF LF_CARE N_DIET N_DIET_WHY
N_DUSUAL N_PRG N_INTK N_EN N_WATER N_PROT N_FAT N_SFA N_MUFA N_PUFA N_N3 N_N6 N_chol
N_tdf N_CA N_PHOS N_NA N_K N_VA N_RETIN N_B1 N_B2 N_NIAC LF_CHIL LF_SAFE LF_S2 LF_S3 LF_S4 LF_S5 LF_S5_1
LF_S6 LF_S7 LF_S8 LF_S9 LF_S10 LF_S11 LF_S12 LF_S13 LF_S14 LF_S15 LF_S16 LF_SECUR LF_SECUR_G

Figure 12: Features selected after filters methods: *t*-test and chi-squared test. This result is used as the input for wrapper method.

Table 4: SVM-based attribute risk factor computation.

Number	Ranked risk factors	Weight
1	Age	0.0890
2	DE1_pr: current status of diabetes	0.0885
3	DN1_dg: kidney failure diagnosis	0.0881
4	BD1_11: (12 years old or older) frequency of drinking for 1 year	0.0873
5	BP8: average sleep time per day	0.0841
6	BP1: usually stressed	0.0370
7	BS3_1: (adult) currently smoking	0.0322
8	N_NA: sodium intake (mg)	0.0321

Table 5: Hypertension-disease rules generated by C4.5.

Number	Rules	Conf.
1	If (age > 70, DE1_pr = 1) - >yes	0.86
2	If (46.5 < age < = 70, BD1_11>=6) - >yes	0.74
3	If (BP8 >8, N_B1 < = 0.88 mg) - >yes	0.73
4	If (marri_1=2, BD1_11>=6) - >yes	0.73
5	If (age <= 46.5, BD1_11<4, N_WATER > 1060.5 ml/day) - >no	0.72
6	If (DN1_dg=1, marri_1=2) - >yes	0.63
7	If (N_B1 < = 0.86 mg/day, N_WATER< = 485.63) - >yes	0.61
8	If (marri_1=1, N_NA >= 3532.6, N_B1 < = 0.86) - >yes	0.60

detection. Another concept that needs to be described is the *k*-factor rule, where *k* is the number of risk factors. The rule here is a three-factor rule.

Based on the key RF, the procedure for the disease-rule generation is shown in Figure 10.

Combined with the disease-domain knowledge that is provided by the domain experts, the TPS has the power to ignore a large amount of the attributes at the very beginning, thereby leaving only the high risk factor support (RFS) attributes [15]. Among these attributes, two attribute sets are

formed for a comparison based on the correlation; that is, if they are very similar to each other, they are strongly correlated, and the TPS will remove the one with the low RFS. Then, a basic association rule-mining algorithm like Apriori and commonly used decision-tree algorithms like C4.5, CART, or Random Forest will be used to generate the k-RF rules. For the first stage of the research, however, only C4.5 is used for the testing data. The algorithmic pseudocode is given in Algorithm 3. First, we calculate the GainRatio for each risk factor r in L, then we choose the factor with the highest GainRatio and create a decision node based on this factor, split the dataset by this node. We repeat these steps until all nodes with appropriate satisfied GainRatio value have been used to generate the tree. The path from the root to leaf is the disease rule.

The reasons for the selection of C4.5 are its simplicity, the accuracy of its results, and the ability to use it on numerical and categorical attributes even with the presence of an overfitting problem.

During the next research stage, algorithms including CART, Random Forest, and KNN will be tested to find the one that fits the most datasets with a high accuracy. Finally, the generated rules will be stored in the HBase in preparation for the next few steps.

4.5. Disease Prediction and Detection. According to the disease rules and the RFS, the highly related key RFS, such as heavy drinking for hypertension, are used to generate the prediction model. This work was completed in the authors' previous study [15]. The multi-RFS will be compared with the disease rule to confirm the patient health condition, and again, this work was completed in [15].

4.6. Cloud-Service Module. The Cloud module plays the roles of data storage and transfer and consists of the public Cloud services GCSql and FCM. Considering the efficiency, connection, and security problems, only commonly used, important, and urgent information like an urgent message from the health-care provider is stored in the GCSql database. Alternatively, the FCM is salient for the communication between the TPS and any relevant device. Again, this work was completed in [15].

5. Experiment

At this stage, the detailed design work of the whole system has been finished, while the implementation work is partially finished. Further, the mobile health-sensor network has been set up in the experimental environment. A large amount of simulated data and a small number of real data that were downloaded from the Korea National Health and Nutrient Examination Survey (KNHANES) [24] were combined with simulated dataset, and the testing data comprises approximately 60,900 patient records, including basic personal information, disease information, and clinical information. The entire cluster has been established, and it can interact with Android devices through the Cloud module. Several devices have been used for the purpose of testing. Meanwhile, an app has been developed for the data collection and a dataset statistical-analysis-result visualization.

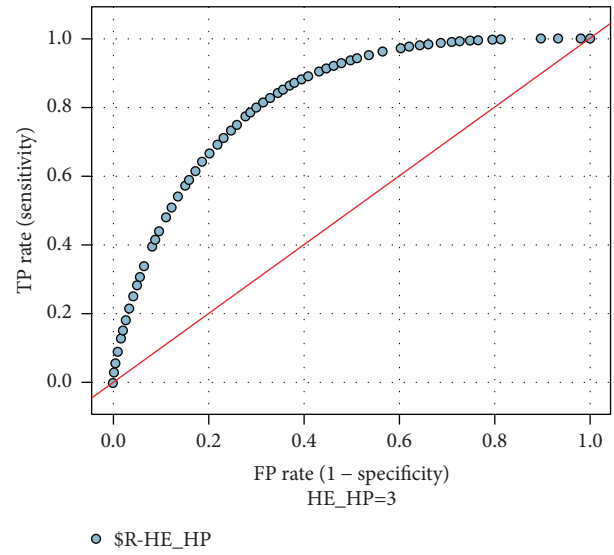


Figure 13: AUC of the trained SVM model. HE_HP = 3.0 is the target attribute.

Table 6: Performance comparison of the two algorithms.

Method	Number of attr.	Sensitivity	Specificity	Accuracy
C4.5	81	0.448	0.692	0.688
SVM	80	0.451	0.705	0.698

A MapReduce-based algorithm called "Disease Count" has been implemented and the pseudocode is given in Algorithm 1. The result of the Disease Count algorithm is given in Table 2. The hypertension data that contains 9383 records has been used as the test dataset, and its statistical-analysis results are given in Table 3. The number of main attributes is 3, but not all of the results are listed in this table due to a space limitation.

The advanced IoT device Fitbit Charge 2 has been used for user sports, sleeping, pulse, and breath detection. The model we used is shown in Figure 11.

Hybrid feature selection mechanism has been used to select highly related features. Among all these attributes, t-test and chi-squared test implemented by using R language have been applied to select the key RFS for hypertension. Finally, 217 features have been selected among all 526 features. The results are given in Figure 12.

SVM-based wrapper feature selection method has been applied to the attributes selected from the previous step. Therefore, 81 features have been selected and the top 8 features are given in Table 4.

For the disease rule generation procedure, the minimum support threshold was set to 0.1, and the minimum confidence threshold was set to 0.3. C4.5 was used to generate the hypertension-disease rules that are shown in Table 5. DI1_dg = 0 means hypertension not diagnosed, while DI1_dg = 1 means hypertension diagnosed. The result is given in Table 5.

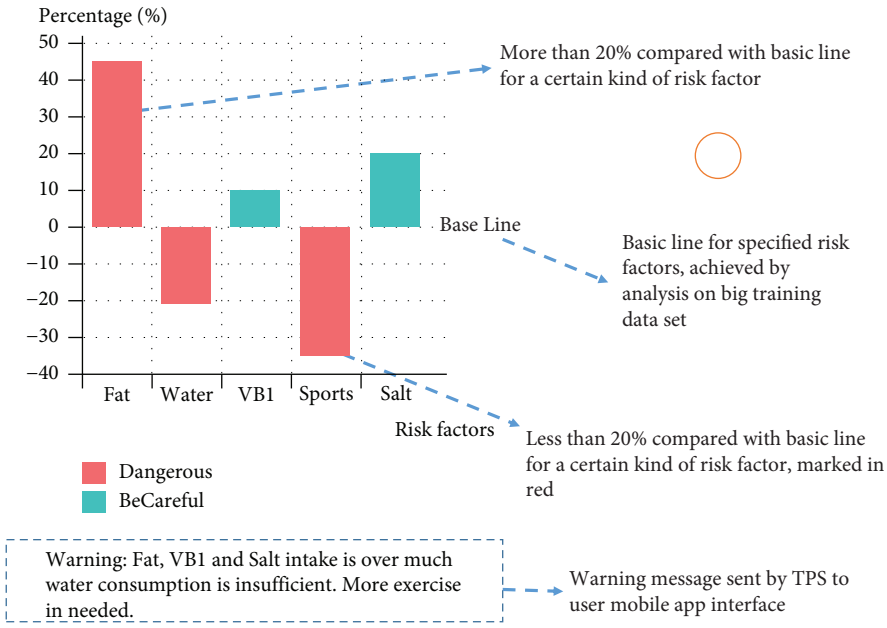


Figure 14: Rule with the risk factor (RF) analysis result visualization. The RFS marked in red mean too much or too few compared with the standard.

For the purpose of the comparison, SVM classification algorithm is also applied to predict the hypertension. We trained one model considered for the variable "HE_HP" for regressor following all the other 80 variables. The support vector machine results were implemented using the "e1071" package in R by performing a 10-folder cross validation using radial basis kernel function which the best parameters are: cost = 1000, gamma = 0.01. The result is given in Figure 13.

We have compared the accuracy of the two methods in terms of sensitivity, specificity, and accuracy. The result is shown in Table 6 below,

From the results, it is possible to draw several conclusions as follows: (1) age and alcohol intake play very important role for hypertension; there is great chance for the elder and heavy drinkers having this disease. (2) The effect of smoking for hypertension is inferior to alcohol. (3) The elders should have light food instead of salty food. (4) SVM performs better result C4.5 in our dataset.

Nevertheless, the accuracy of both algorithms is not satisfactory. This is due to the challenges of collecting big healthcare data, and this issue has been depicted in introduction session. For the next stage of research, we will focus on solving this problem.

The analysis results are directly and visually displayed in the user devices. The interpretations of the analysis results are shown in Figure 14. The authors have published another paper [15] wherein a simple disease-rule visualization method is discussed, since it is also a challenging work.

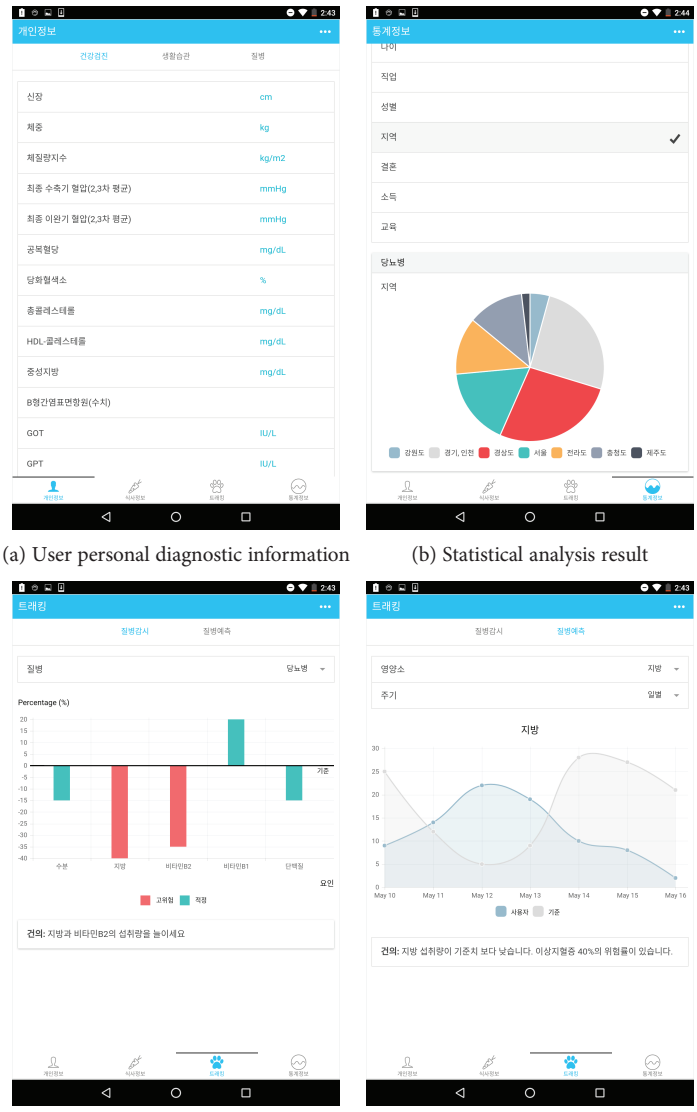
Figure 14 illustrates the disease-detection-result visualization interface of the designed app of this study. The x-axis of the coordinate lists rank the key RFS of a certain kind of disease, and it is also the basic line that is based on the training big data analysis result (e.g., a standard factor like nutritional intake will be visible for a healthy patient, but the concrete value will be hidden from the figure). The y-axis is the percentage of the intake that exceeds or is inferior to the standard factor intake; the disease rules consist of these factors. For a certain disease, there is usually more than one rule (consisting of RFS) that is related to the disease. Compared with these rules, if the matching rate $> \beta$ (expert-defined threshold, e.g., 80%), the system will treat this patient segment as the disease holder.

The figures below are GUI of our system, which is implemented based on Ionic using hybrid programming techniques. The characters are in Korean since it is mainly developed for Korean users so far. Selected user interfaces are given in Figure 15.

6. Conclusion and Future Work

In the present work, Hadoop, Spark, and DM techniques are exploited to design a comprehensive, real-time, and intelligent mobile health-care system that can facilitate a step-by-step process for disease detection and prediction. The purpose of this work is the provision of a practical assistant system for self-based user health care, as well as the design of a complementary system for patient disease diagnosis. During the experiment section, firstly, disease data stored in distributed environment has been retrieved by MapReduce method and analysed by statistical method to give us an overview of the data. Then both statistical methods and DM methods are used to select the features related to hypertension disease. These attributes are the risk factors as well. Based on these factors, C4.5 and SVM methods are used to generate the classifier model for disease prediction. Finally, we displayed the analysis result on users' mobile devices.

An overview and a guide are described in detail for a future work as well. For the next stage of research, after the implementation of the whole system, in-depth simulations will be performed to validate the systemic performance in



(a) User personal diagnostic information

(b) Statistical analysis result

(c) Disease detection interface

(d) Disease prediction interface

Figure 15: User interfaces of the system.

terms of the application of the proposed system in a real environment. The TPS and algorithms like C5, Random Forest, and other algorithms will be run on the TPS cluster to compare the efficiency and the accuracy. The procedure for the disease detection and prediction will be optimized continuously and extended to other chronic diseases. Finally, it is hoped that this system will contribute to health-care academia as well as the industry.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (Grant no.2017R1A2B4010826) and the MSIP (Ministry of Science, ICT and Future Planning), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2017-2013-0-00881), supervised by the IITP (Institute for Information & Communication Technology Promotion) and also supported by the National Natural Science Foundation of China (61702324).

References

[1] A. V. Chobanian, G. L. Bakris, H. R. Black et al., "Seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure," *Hypertension*, vol. 42, no. 6, pp. 1206–1252, 2003.

[2] P. D. Kaur and I. Chana, "Cloud based intelligent system for delivering health care as a service," *Computer Methods and Programs in Biomedicine*, vol. 113, no. 1, pp. 346–359, 2014.

[3] H. Horiguchi, H. Yasunaga, H. Hashimoto, and K. Ohe, "A user-friendly tool to transform large scale administrative data into wide table format using a MapReduce program with a

pig Latin based script," *BMC Medical Informatics and Decision Making*, vol. 12, no. 1, p. 151, 2012.

[4] R. C. Taylor, "An overview of the Hadoop/MapReduce/HBase framework and its current applications in bioinformatics," *BMC Bioinformatics*, vol. 11, article S1, Supplement 12, 2010.

[5] M. C. Schatz, "CloudBurst: highly sensitive read mapping with MapReduce," *Bioinformatics*, vol. 25, no. 11, pp. 1363–1369, 2009.

[6] V. Sharma, K. Mankodiya, F. De La Torre et al., "SPARK: personalized Parkinson disease interventions through synergy between a smartphone and a smartwatch," in *Design, User Experience, and Usability. User Experience Design for Everyday Life Applications and Services. DUXU 2014*, A. Marcus, Ed., vol. 8519 of Lecture Notes in Computer Science, pp. 103–114, Springer, Cham, 2014.

[7] H. Kim, M. I. M. Ishag, M. Piao, T. Kwon, and K. H. Ryu, "A data mining approach for cardiovascular disease diagnosis using heart rate variability and images of carotid arteries," *Symmetry*, vol. 8, no. 6, p. 47, 2016.

[8] S. Amendola, R. Lodato, S. Manzari, C. Occhiuzzi, and G. Marrocco, "RFID technology for IoT-based personal healthcare in smart spaces," *IEEE Internet of Things Journal*, vol. 1, no. 2, pp. 144–152, 2014.

[9] M. J. Huang, M. Y. Chen, and S. C. Lee, "Integrating data mining with case-based reasoning for chronic diseases prognosis and diagnosis," *Expert Systems with Applications*, vol. 32, no. 3, pp. 856–867, 2007.

[10] S. H. Ha and S. H. Joo, "A hybrid data mining method for the medical classification of chest pain," *International Journal of Computer and Information Engineering*, vol. 4, no. 1, pp. 33–38, 2010.

[11] H. J. Aboumatar, K. A. Carson, M. C. Beach, D. L. Roter, and L. A. Cooper, "The impact of health literacy on desire for participation in healthcare, medical visit communication, and patient reported outcomes among patients with hypertension," *Journal of General Internal Medicine*, vol. 28, no. 11, pp. 1469–1476, 2013.

[12] Y. Piao, M. Piao, C. H. Jin et al., "A new ensemble method with feature space partitioning for high-dimensional data classification," *Mathematical Problems in Engineering*, vol. 2015, Article ID 590678, 12 pages, 2015.

[13] M. Herland, T. M. Khoshgoftaar, and R. Wald, "Survey of clinical data mining applications on big data in health informatics," in *2013 12th International Conference on Machine Learning and Applications*, vol. 2, pp. 465–472, Miami, FL, USA, December 2013.

[14] R. Fang, S. Pouyanfar, Y. Yang, S. C. Chen, and S. S. Iyengar, "Computational health informatics in the big data age: a survey," *ACM Computing Surveys (CSUR)*, vol. 49, no. 1, article 12, 2016.

[15] D. Li, H. Park, E. Batbaatar, Y. Piao, and K. Ryu, "Design and partial implementation of health care system for disease detection and behavior analysis by using DM techniques," in *2016 IEEE 14th Intl Conf on Dependable, Autonomic and Secure Computing, 14th Intl Conf on Pervasive Intelligence and Computing, 2nd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech)*, pp. 781–786, Auckland, New Zealand, August 2016.

[16] "Welcome to Apache™ Hadoop," http://hadoop.apache.org/.

[17] "Apache™ Spark, a fast and general engine for large-scale data processing," http://spark.apache.org/.

[18] J. Han, J. Pei, and M. Kamber, *Chapter 1 Introduction Data Mining: Concepts and Techniques*, Elsevier, Waltham, MA, USA, 2011.

[19] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[20] H. H. Hsu, C. W. Hsieh, and M. D. Lu, "Hybrid feature selection by combining filters and wrappers," *Expert Systems with Applications*, vol. 38, no. 7, pp. 8144–8150, 2011.

[21] "Apache Sqoop™," http://sqoop.apache.org/.

[22] B. Lee and E. Jeong, "A design of a patient-customized healthcare system based on the hadoop with text mining (PHSHT) for an efficient disease management and prediction," *International Journal of Software Engineering and Its Applications*, vol. 8, no. 8, pp. 131–150, 2014.

[23] "Apache Spark™," http://spark.apache.org/streaming/.

[24] "KNHANES," https://knhanes.cdc.go.kr/knhanes/sub03/sub03_02_02.do.

Hindawi
Journal of Sensors
Volume 2018, Article ID 6464036, 11 pages
https://doi.org/10.1155/2018/6464036

Research Article

IoT Smart Home Adoption: The Importance of Proper Level Automation

Heetae Yang,[1] Wonji Lee,[2] and Hwansoo Lee[3]

[1]*Science and Technology Policy Institute, Sejong National Research Complex, 370 Sicheong-daero, Sejong-si 30147, Republic of Korea*
[2]*Interdisciplinary Graduate Program in IT Law, Dankook University, 152 Jukjeon-ro, Suji-gu, Yongin-si, Gyeonggi-do 16890, Republic of Korea*
[3]*Department of Convergence Security, Dankook University, 152 Jukjeon-ro, Suji-gu, Yongin-si, Gyeonggi-do 16890, Republic of Korea*

Correspondence should be addressed to Hwansoo Lee; hanslee992@gmail.com

Received 13 November 2017; Accepted 9 May 2018; Published 22 May 2018

Academic Editor: Mucheol Kim

The word "smart" has been used in various fields and is widely accepted to mean intelligence. Smart home service, one of the representative emerging technologies in the IoT era, has changed house equipment into being more intelligent, remote controllable, and interconnected. However, the intelligence and controllability of a smart home service are contradictory concepts, under certain aspects. In addition, the level of intelligence or controllability of a smart home service that users want may differ according to the user. As potential users of smart home services have diversified in recent years, providing the appropriate functions and features is critical to the diffusion of the service. Thus, this study examines the smart home service features that current users require and empirically evaluates the relationship between the critical factors and the adoption behavior with 216 samples from Korea. The moderating effect of personal characteristics on behavior is also tested. The results of the analysis provide various theoretical and practical implications.

1. Introduction

We are experiencing a new era of Internet of Things (IoT), where many electronic devices surrounding us are interconnected by a network [1]. This paradigm enables copious amounts of data to be stored, processed, and conferred in a proficiently interpretable form without human invention. The emerging of IoT also sheds new light on the concept of a "smart home." IoT-enabled house equipment allows for a smart home to be more intelligent, remote controllable, and interconnected.

The global smart home market is expected to grow to USD 119.26 billion by 2022 [2]. Global companies (e.g., Google, Amazon, and Samsung Electronics) are entering this huge market, and they are providing innovative services and products to take advantage of the growing market. Many start-ups are also making efforts to join this growing market. The smart home has been drawing attention recently due to the IoT, but it is not a new concept. In fact, the concept of a smart home has been discussed since 1980, and it has evolved from traditional home automation (so-called networked homes, ubiquitous homes, and intelligent and interactive homes).

Despite the long history and growing interest, smart home service has not been widely accepted. There are many reasons (e.g., high device prices, limited consumer demand, and long device replacement cycles) preventing smart home diffusion. The largest barrier is due to a lack of technology to establish the infrastructure of a smart home [3]. Edwards and Grinter [4] argued that another reason is that the issues and social aspects of adoption and diffusion of a smart home service have been overlooked. Prior studies on smart homes have been conducted without reflecting user characteristics or their environments. Most approaches are based on experiments or solely focused on the technology. Technological or engineering perspectives on smart homes have failed to interpret potential users' actual needs from a smart home.

Thus, this research aims to find what users really want from smart home services and investigate which features affect a user's intention to leverage the service. This study will

Table 1: Evolution of smart home services.

Year	Phase	Technical background	Main function
1990s	Home automation	Broadband Internet	Household automation
2000s	Home network	Smart phone and app	Remote monitoring & control
2010s	Smart home	IoT & AI	Context awareness

define the user acceptance factors from a new perspective and present a theoretical model to verify precedent factors and outcomes. Through this empirical and behavioral analysis, it will determine the concept of "smart" as reflecting a user's actual smart home needs and its major functional features. In order to confirm that the needs for smartization sought by people vary depending on their individual characteristics and environment, the research also studies how aspects including the type of housing, gender, age, and prior experience affect user intentions.

2. Related Background

2.1. Smart Home. A smart home refers to a residence equipped with a communication network, high-tech household devices, appliances, and sensors that can be remotely accessed, monitored, and controlled and that provide services responding to the residents' needs [5]. Although the widespread diffusion of high-speed Internet in the late 1990s provided the opportunity for the home network business to grow, it was not until the late 2000s that smart homes began to be installed, which is when smart phones were popularized. Initially, a smart home was defined using various names, such as a home network, a digital home, home automation, and an intelligent home. In the mid-2010s, it has been leaning towards a combination of Internet of Things (IoT) and a situation-aware smart home (Table 1).

A smart home is an advanced form of traditional home automation. An early definition of a smart home, which was influenced by home automation, is using common communication devices to integrate with a variety of services at home, assuring economic, secure, and comfortable operation of the home [6]. Thus, smart home service was used to manage environmental systems like lighting and heating [3]. These days due to technological development, a smart home service monitors user activities and the internal environment at home (Figure 1). Moreover, a smart home provides services that fulfill the demands and needs of a user.

Recently, smart home services are evolving as they approach AI. The intelligent personal assistant "Alexa," developed by Amazon Lab126, has been installed in a wide range of products. LG Electronics has adopted Alexa throughout its smart home product line. For example, if a user calls "Alexa" from a smart refrigerator, the user can access services such as searching news, online shopping, and checking schedules. In addition, China smart home manufacturer Xiaomi is planning to target the smart home market as part of its long-term vision. Xiaomi launched an air purifier that can be remotely controlled by a smart phone and developed a smart module that can be inserted into all appliances such as refrigerators, air conditioners, and washing machines. Apple is developing an AI speaker that supports "Apple HomeKit," expected to provide voice support as a hub to control home kit products. Thus, smart home services are developing and proliferating by adopting IoT and AI.

Prior studies on smart homes are based on a technical or a partial approach. For instance, [7] suggested a smart floor technology for a smart home which examines who, where, and what a user is doing. Adami et al. [8] proposed a wrist gadget that monitors users' habits at home. Andoh et al. [9] suggested a biometrics monitoring system for analyzing the pulse and respiratory rate at home. Koskela and Väänänen-Vainio-Mattila [10] conducted experiments with only young nontechnical professionals to determine whether users welcome smartness through familiar communication devices such as PCs, TVs, and mobile phones. Paetz et al. [11] proposed an automated energy management system to test residents. These studies have contributed greatly to improve the completeness of a smart home service. However, as research on business and user perspectives for market revitalization was lacking, more studies are requisite to encourage the proliferation of smart home services.

2.2. Smart Home Service Adoption. The word "smart" has been used in various fields such as smart phones, smart TVs, and smart learning, including smart homes. Although it has a slightly different meaning in each concept, it generally means "intelligent," which can be interpreted as the concept of a weak level of artificial intelligence (AI). However, whether such a concept represents an intelligence service that can perfectly substitute the decision-making process of human beings requires debate. Humans in general will be reluctant to delegate all of the decision-making authority to machines considering their search for freedom, uncertainty, and distrust in technology. In addition, the level of smartness people demand would also vary depending on their individual characteristics and environment. Some people have vague fears about intelligent and smart things. For instance, when AlphaGo beat a human in the Go game, some people had negative perspectives on AI because a computer can control or be detrimental to people. Thus, the "smart" that people want may entail a limited scope of intelligence which is under the control of human beings, unlike the theoretical point of view.

Smart home service acceptance research has been active since the mid-2000s. Most studies have extended the technology acceptance model (TAM) or the unified theory of acceptance and use of technology (UTAUT) and have focused on specific groups such as the elderly, the disabled, and patients. Leeraphong et al. [12] demonstrated that self-efficacy plays an important role in the

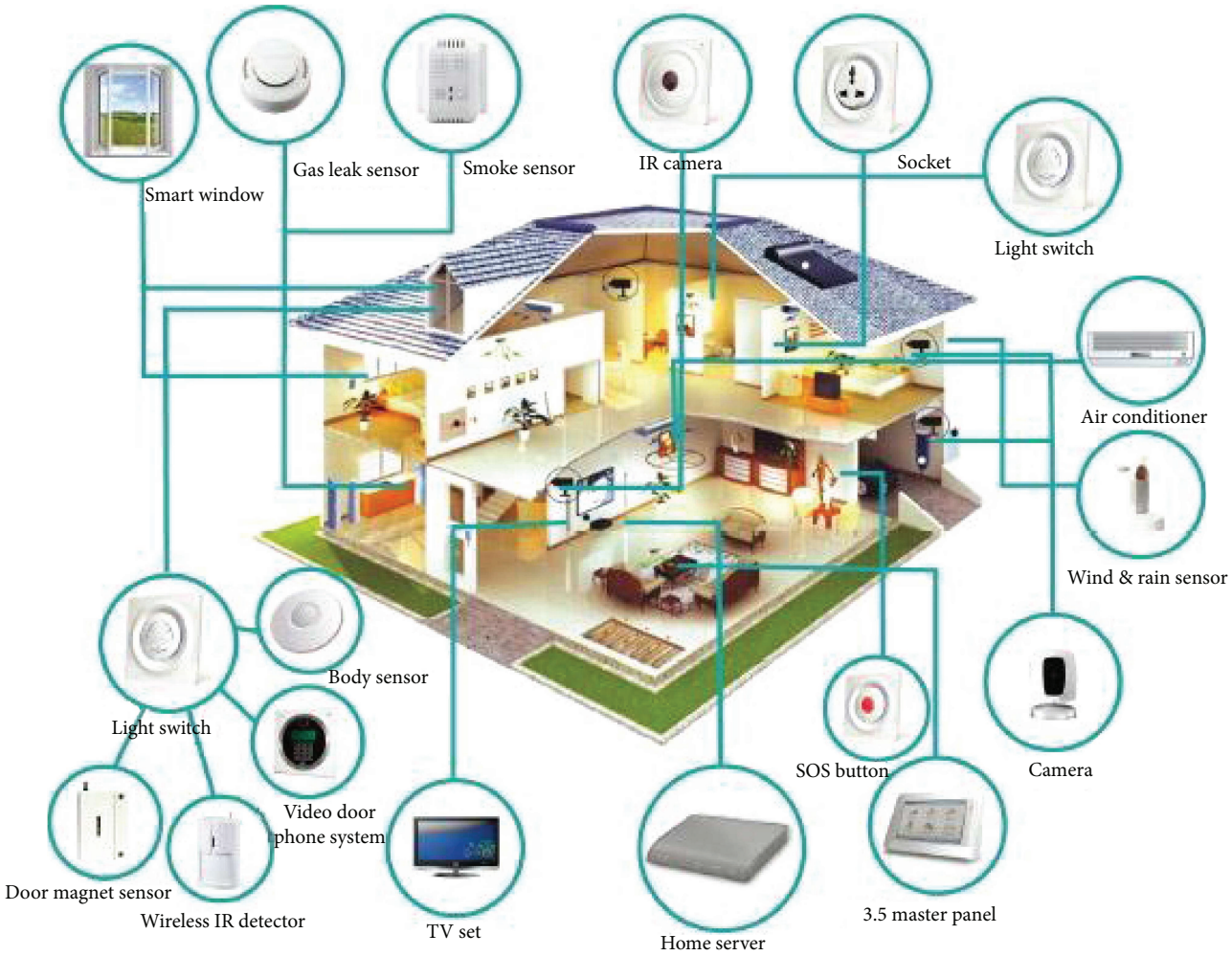


Figure 1: Smart Home (http://smarthomeenergy.co.uk).

acceptance of a smart home for the elderly. In the Alaiad and Zhou study [13], patients' expectations of quality of life were found to play a key role in smart home acceptance. While expanding on the existing theories, attempts have been made to search for new factors. Fan et al. [14] mentioned the continued development of AAL (ambient assisted living) which assists disabled and elderly people especially with chronic diseases, to enhance their well-being and enable independent living. Vadillo et al. [15] conducted research on telecare system adoption, which is one type of smart home service, and found perceived usefulness is important for the intention to use the system.

Recently, as smart services become more common, research on the acceptance of general users has become active. Bao et al. [16] examined the mobile smart home environment and found that compatibility and safety are key factors affecting mobile smart home acceptance. Park et al. [17] foresee that the future concept of smart home acceptance with IoT technologies will expand by presenting accessibility to wireless networks, as well as the compatibility of various operation systems, languages, and frameworks. Furthermore, this study empirically proved that enjoyment, compatibility, connectedness, control, and cost can motivate the acceptance of a smart home. Ji et al. [18] pointed out that controllability is significant in adopting a smart home. In addition, intelligent control is able to learn user behavior and demand automatically and provide a smarter control of appliances.

Some studies emphasize the importance of security and privacy in smart home acceptance. For instance, Tanwar et al. [19] identified that the level of security in households is distinctively conceivable in accepting a smart home service. Elmasllari and Al-Akkad [20] argued that clear privacy and data protection is needed to operate smart energy systems. La Marra et al. [21] stated that security and safety risks which affect user behavior can also be mitigated by adding configurable systems for security policy enforcement.

However, as the existing discussions are abstract, a fundamental understanding is necessary to characterize smart home services. In addition, past studies have conducted an empirical analysis on a specific group, but the popularization of smart home services now requires more general discussions for diverse user classes.

2.3. Critical Factors for Smart Home Service. The initial smart home service was promoted through the automation of the

domestic system, aiming for convenience, comfort, stability, amenity, health, reduction of household labor, and energy efficiency. Since then, developments of the wireless Internet and smart phones have extended the concept of a smart home to services that can be remotely controlled anytime and anywhere. In the IoT era, household electrical appliances and information and communication devices are interconnected, and the smart home is developing into a form of an artificial intelligence service that operates by self-understanding the behaviors of the residents. Therefore, the smart home in the IoT era is a concept that adds interconnectedness to the traditional characteristics of automation and remote controllability [22]. Service stability, security, and privacy also have been suggested as important factors that may hinder user acceptance [23]. These factors can be summarized as the reliability of the service. The smart home environment is a factor that must be considered because it is closely related to the user's life and can cause serious damage in the case of a dangerous situation. Thus, automation, remote controllability, interconnectedness, and reliability can be summarized as crucial factors for accepting a smart home service.

Automation is defined as the "execution by a machine agent (usually a computer) of a function that was previously carried out by human" [24]. Home automation was the initial name of the smart home service, and the automation of households and home infrastructure was a key goal of the early smart home. Automation has become prominent in recent years because it has become more affordable and simple through the development of information technology. In recent years, an interest in AI has enabled higher-level automation. AI technology can advance the function of a smart home by assisting users in an intelligent way [25]. Hence, technology is one of the critical features of a smart home.

The virtue of a smart home is that it can be controlled remotely by mobile devices. This is a core feature of a smart home system since users prefer to instantly control smart home services such as controlling lamps, curtains, and information appliances [10]. However, to design an intelligent and remote-controllable smart home system, a network connection is essential. Many networks exist with a variety of features such as Bluetooth IEEE 802.15.4, Z-Wave, and Wi-Fi. To enable remote control, networks should be standardized and interconnected to expand the use of smart home services. Most electronic devices support the Wi-Fi protocol, which allows home devices to be controlled by mobile devices. When remote control is possible, the general concept of smart, anywhere and anytime, can actually be implemented.

Interconnectedness is defined as the ability of devices, applications, and services to be connected with each other to work together [26]. To proliferate a smart home, devices should able to adapt to changes in the preferences, requirements, and needs of a user [4]. The system should easily connect to new devices in a smart home. It is critical to correspond through the network for a smart home to function properly. However, many types of network and communication protocols are a barrier in reality [5]. Networks can be wired or wireless, and other types of communications exist. At present, the technical standard is inadequate due to the high cost of satellite links and the limited transmission between electronic devices [27].

Technical errors in integrated smart homes can be a concern to potential users. The reliability of smart home services depends not only on the fact that the technology will not malfunction but also on the fact that the technological components will function flawlessly while providing an accurate service [5]. Users' trust in service providers is an important issue for the diffusion of smart home services. Currently, technologies for smart homes are limited to predicting human behavior. Smart home services must be able to provide reliable measurements and algorithms to assess a user's vital signs or lifestyle [27].

3. Research Model and Hypotheses Development

The research model of this study was developed as shown in Figure 2, with the following essential antecedent factors that affect a user's intentions for smart home services: automation, controllability, interconnectedness, and reliability.

3.1. Perceived Automation. Automation is a term referring to the automation of housework and household activities such as the control of lighting, heating systems, and ventilation. This kind of automation enables users to be comfortable, live conveniently, be secure, and be energy efficient. In addition, it monitors elderly and disabled people to ensure suitable care [28]. The acceptance of automation has increased in smart homes recently due to the enhanced affordability and simplicity through upgraded technology [29]. Sági et al. [28] proposed a controller to optimize power management to reduce the energy consumption of a smart home automation system. Luor et al. [29] authenticated the correlation between the user attitude and efficient smart home automation function.

H1: perceived automation is positively associated with smart home adoption intention.

3.2. Perceived Controllability. Controllability is the ability to do whatever a user needs with the given system that is under control [30]. Liu et al. [31] confirmed that control is a critical issue in most complex networks. Users remotely control smart home systems by accessing services on smartphones, mobile phones, and computers [32, 33]. Due to the increased number of sensors and multitouch screens, mobile devices became a pivotal user interface in smart home systems [34]. In addition, Roduner et al. [35] implicated that user interfaces on mobile devices are key for controlling a smart home system.

H2: perceived controllability is positively associated with smart home adoption intention.

3.3. Perceived Interconnectedness. Interconnectedness is defined as the ability to work together reliably owing to the fact that a discrete manufacturer exists [26]. Many studies have revealed the importance of interconnectedness while adopting new IT services. The study by [36] defined that

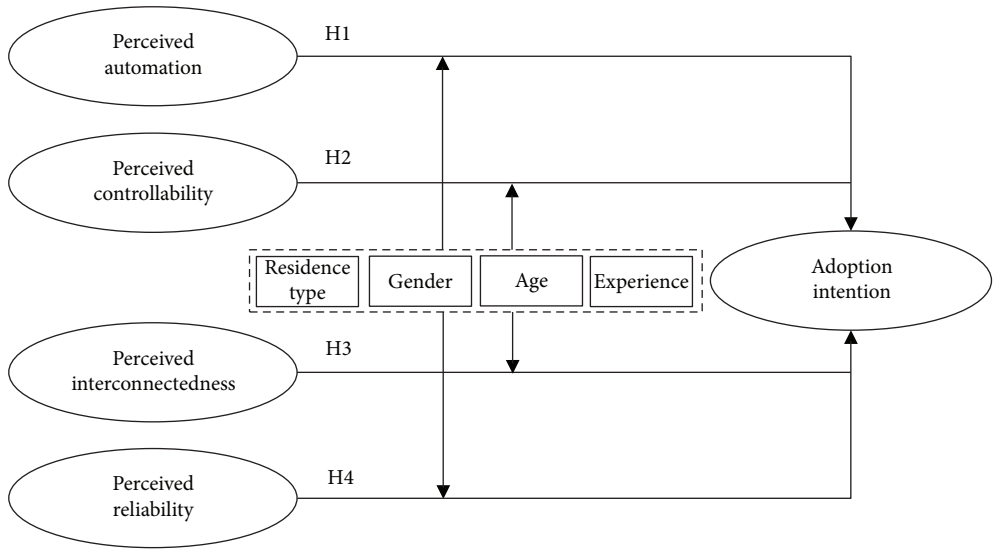


Figure 2: Research model.

compatibility was a key component in perceiving the efficiency of mobile multimedia services. Yang et al. [37] pointed out that technical compatibility affects users' perceived usefulness of wearable devices. IoT applications or services cannot be said to be smart home simply because other devices monitor or adjust one home application, sensor, or device. Since the interworking of multiple devices and a central device is essential, a smart home system requires a platform to provide a common framework [27].

H3: perceived interconnectedness is positively associated with smart home adoption intention.

3.4. Perceived Reliability. Reliability between a manufacturer and a user is an important factor in user behavior. In the Maslow theory of human motivation safety, security and protection are the second needs to be satisfied after fulfilling basic physiological needs like food, water, and shelter [38]. Services like alarms and at-home security should be operating properly. Keen et al. [39] suggested that e-commerce represents a strategic indication of trust for consumer-marketer relationships. Chan et al. [27] indicated the reliability of a sensor and data processing system as a one of main determining factors of adopting a smart home. Friedewald et al. [40] verified that a smart home system should aim to be dependable which leads to successful innovation from the socioeconomic and technological constellation.

H4: perceived reliability is positively associated with smart home adoption intention.

Smart home services are no longer a specialized service for a specific group of people such as housewives, patients, or the elderly but are now developing into a more public service that the general public can use for a more convenient lifestyle. Such trends call for the need of a more detailed analysis of the actual motivations of diverse groups of users in order to generalize the research results and render them a reflection of the current times. As there could be a difference in the relationship between variables representing conventional users and new groups

Table 2: Characteristics of the respondents.

Characteristics	Respondents ($n = 216$)	
	Number	Percentage
Gender		
Male	111	51.4
Female	105	48.6
Age		
16–19	28	13.0
20–29	37	17.1
30–39	45	20.8
40–49	46	21.2
50–59	38	17.5
60+	22	10.4
Education		
Less than high school	44	20.4
College or university	150	69.4
Advanced degree	22	10.2
Occupation		
Official worker	60	27.8
Service worker	8	3.7
Professional/researcher	12	5.5
Self-employer	14	6.5
Public service worker	3	1.4
Student	66	30.6
Housewife	34	15.7
Other	19	8.8
Residence type		
Apartment/multifamily house	114	52.8
Single house	102	47.2
Experience of smart home service		
Yes	46	21.3
No	170	78.7

of users when considering their characteristics, the research investigated the differences between groups according to the type of housing, gender, age, and use experience, as moderator variables.

4. Research Method

4.1. Data. This study conducted an online survey in October 2015 to evaluate the research model. Data was collected by a professional survey company in Korea. The final 216 collected samples were used for analysis. The demographic distribution of the samples is as follows: 111 males (51.4%), 105 females (48.6%), 10s (28, 13%), 20s (37, 17.1%), 30s (45, 20.8%), 40s (46, 21.2%), 50s (38, 17.5%), and over 60 (22, 10.4%). The distribution of the samples was balanced for analysis (Table 2).

4.2. Instrument Development. All measurement items to measure latent constructs were developed based on previous studies. Responses were collected based on a 5-point Likert scale. The items of each construct are shown in Table 3. The PLS-SEM (partial least squares structural equation modeling) analysis method and the Smart PLS 2.0 tool were used to verify the research hypotheses after testing the convergent validity of each construct and discriminant validity. Furthermore, in order to compare the route coefficients according to the residence type, sex, age, and experience, we use the following formula given by Chin and Dibbern [41].

$$t = \frac{\text{Path}_{\text{sample1}} - \text{Path}_{\text{sample2}}}{\sqrt{\left((m-1)^2/(m+n-2)\right) \times \text{s.e.}^2{}_{\text{sample1}} + \left((n-1)^2/(m+n-2)\right) \times \text{s.e.}^2{}_{\text{sample2}}} \times \sqrt{(1/m)+(1/n)}} \sim t_{m+n-2} \quad (1)$$

where, $\text{Path}_{\text{sample1 or sample2}}$ is the path coefficient in the subsample 1 or 2, m is the number of cases in subsample 1, n is the number of cases in subsample 2, and $\text{s.e.}_{\text{sample1 or sample2}}$ is the standard error of the path coefficient in subsample 1 or 2.

5. Data Analysis and Results

5.1. Measurement Model. The reliability and validity of the constructs were checked. The reliability of latent variables can be confirmed by Cronbach's α and composite reliability (CR). Generally, reliability is considered to be satisfied when Cronbach's α exceeds 0.7 [49]. Even though Cronbach's α for perceived controllability and perceived reliability are lower than 0.7, the CR value was sufficiently big, so there was no significant problem in reliability (Table 4). Confirmatory factor analysis was conducted to test the convergent validity of each construct. In Table 4, all factor loadings exceeded 0.6, the minimum required to assure convergent validity of the constructs [50], and the AVE for each construct exceeded 0.50 [51]. As a result, convergent validity is established. Discriminant validity was demonstrated by confirming that the square root of the average variance extracted (AVE) for each construct is higher than the corresponding interconstruct correlations. There was no critical issue in cross-loading for the discriminant validity (see Tables 5 and 6).

5.2. Hypotheses Testing. Structural equation modelling results are presented in Figure 3, and the hypothesis tests are summarized in Table 4. A bootstrapping resampling technique was employed to calculate the corresponding t-values for each hypothesized relationship. As summarized in Figure 3, in the analysis of the whole sample, out of the 4 hypotheses, three were supported. Perceived controllability, perceived interconnectedness, and perceived reliability were significant factors influencing adoption intention, supporting H2, H3, and H4, and explaining 47.9% of the variance (H2: $\beta = 0.327$, t-value = 3.170, and $p < 0.001$; H3: $\beta = 0.232$, t-value = 2.363, and $p < 0.05$; and H4: $\beta = 0.154$, t-value = 2.642, and $p < 0.01$; $R^2 = 0.479$). However, the path from perceived automation to adoption intention was insignificant, rejecting H1.

However, the results were very diverse when analyzed by grouping according to the residence type, gender, age, and experience (Table 7). For example, in the analysis for general home residents, H1, H2, and H4 were supported but H3 was not. On the other hand, only H2 and H3 were supported for apartment residents. In terms of gender, H2 and H3 were supported in males but H2 and H4 were supported in females. When analyzed by age, H2 and H3 were supported by those under the age of 39 but H2 was supported only by those over 40 years of age. H1 and H4 were supported for current smart home service users, and H2 and H3 were supported for potential users.

6. Discussion

6.1. Findings. The objective of this study was to understand and explain customers' behavioral intentions to adopt smart home services. Unlike the previous research that examined the user behavior associated with smart home service adoption based on acceptance theories, this study captures the characteristics of smart home services and presents a new theory and model. The empirical analysis of the proposed research model demonstrated various implications.

In all samples, three factors: controllability, interconnectedness, and reliability had a significant impact on the acceptance behavior of a smart home service. It is very interesting that automation does not have a significant effect. This can be interpreted as follows: people generally seek relatively safer

Table 3: Survey items.

Construct	Item number	Measurement items	References
Perceived automation	PA1	Smart home services help the residents proactively without human intervention.	[29, 42]
	PA2	Smart home services provide autoadjusted control.	
Perceived controllability	PC1	I can control every electrical device of smart home services through simple operation.	[43, 44]
	PC2	It is convenient to control smart home services anywhere at any time.	
Perceived interconnectedness	PI1	Smart home devices are interconnected with each other.	[45, 46]
	PI2	Smart home services by integrating different device venders do not create problems.	
Perceived reliability	PR1	I am not worried to use smart home services because other people or organizations may be able to access my account.	[43, 47]
	PR2	There will not be much potential loss associated with disclosing personal information to the smart home service provider.	
	PR3	I think smart home service providers are reliable.	
Adoption intention	AI1	Using smart home services is worthwhile.	[48]
	AI2	I intend to use smart home services in the future.	
	AI3	I predict I would use smart home services in the future.	

Table 4: Validity of constructs.

Construct	Items	Factor loading	t-value	AVE (>0.5)	Composite reliability (>0.6)	Cronbach's alpha (>0.7)
Perceived automation	PA1	0.947	97.763	0.902	0.965	0.946
	PA2	0.954	119.742			
Perceived controllability	PC1	0.898	61.407	0.903	0.949	0.893
	PC2	0.831	23.021			
Perceived interconnectedness	PI1	0.935	94.090	0.748	0.855	0.667
	PI2	0.902	38.725			
Perceived reliability	PR1	0.714	7.651	0.843	0.915	0.815
	PR2	0.769	9.760			
	PR3	0.769	12.624			
Adoption intention	CONT1	0.947	114.355	0.564	0.795	0.647
	CONT2	0.964	163.208			
	CONT3	0.938	101.322			

Table 5: Correlations of the constructs and square root of AVE.

	PA	PC	PI	PR	AI
PA	**0.90**				
PC	0.68	**0.86**			
PI	0.54	0.78	**0.92**		
PR	0.36	0.56	0.51	**0.75**	
AI	0.50	0.66	0.62	0.49	**0.95**

Table 6: Construct cross-loadings.

	PA	PC	PI	PR	AI
PA1	**0.947**	0.637	0.492	0.355	0.457
PA2	**0.954**	0.650	0.540	0.334	0.486
PC1	0.716	**0.898**	0.615	0.513	0.629
PC2	0.426	**0.831**	0.757	0.448	0.497
PI1	0.543	0.735	**0.935**	0.480	0.616
PI2	0.448	0.696	**0.902**	0.453	0.507
PR1	0.162	0.321	0.303	**0.714**	0.249
PR2	0.188	0.285	0.301	**0.769**	0.308
PR3	0.389	0.564	0.482	**0.769**	0.472
AI1	0.446	0.581	0.553	0.437	**0.947**
AI2	0.482	0.662	0.598	0.463	**0.964**
AI3	0.484	0.625	0.602	0.490	**0.938**

and more effective remote management features rather than highly advanced automated services. People may want the devices of a smart home to be under their control rather than being fully automated because a home is safe and represents their personal space where they can rest. Considering that controllability is the most important antecedent for adoption, it becomes apparent that the automation that people

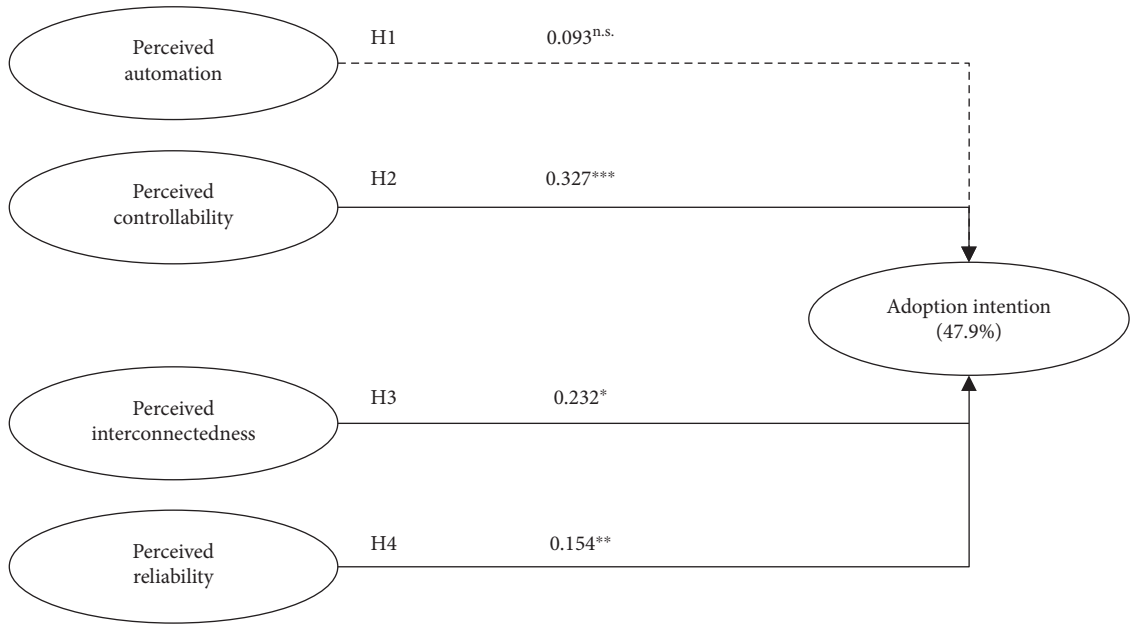


Figure 3: Results of the structural model: $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$ (two tailed).

Table 7: Result of path coefficients by grouping.

	Residence type (*N*)		Gender		Age		Experience	
	APT (114)	House (102)	M (111)	F (105)	<39 (110)	40+ (106)	Yes (46)	No (170)
H1	−0.053	0.271***	0.153	0.004	0.015	0.159	0.332**	0.047
H2	0.367*	0.317*	0.291*	0.378**	0.256*	0.378**	0.264	0.310*
H3	0.301*	0.129	0.251*	0.217	0.337*	0.134	0.062	0.272*
H4	0.148	0.165*	0.142	0.184*	0.157	0.177*	0.342**	0.140

$^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$.

want is eventually intelligent and represents an optimal controllability that is close to a limited form of automation.

There are also various findings in group comparison analysis. Controllability and interconnectedness significantly affect adoption for people living in apartments while automation and reliability are considered to be important for general home residents. The difference in the level of infrastructure between apartments and general homes may be an influence. Recently, new apartments in Korea have been provided with smart home services (remote heating management, gas shutdown, etc.). As a result, apartment residents seem to want more extensive and precise control. As general homes do not provide any networked and automated functions to control households, general home residents may want the automation and reliability of a smart home service.

When comparing gender, men emphasize interconnectivity while women emphasize reliability. Similar results were observed in the age-related comparisons. This can be attributed to the risk avoiding tendency of women and older people. It is understandable that they prefer these factors because they seek stability compared to men or young people. Men and young people tend to prefer interesting and innovative services, and the interconnectedness of the smart home can serve as a factor in meeting these needs.

Those who have experienced similar services emphasize automation and reliability. In fact, it is likely that those who do not place much significance on control and interconnection have experienced that the control and connectivity of past services did not guarantee usability.

6.2. Contributions. This study makes several contributions to theory. First, it presents the specific success factors of smart home service adoption and empirically analyzes the relationship with the acceptance behavior. Most studies have presented abstract success factors (e.g., usefulness and enjoyment) based on technology acceptance theory, but this study derived the detailed critical factors through a literature review. Second, this study captures the concept of the word "smart" in the smart home service. According to the result, the smart that people desire is close to intelligent control but not fully automated. This explains why past services (e.g., home automation and the networked home) that are similar to a smart home have not spread. The past services have not been able to meet the desired level of automation for users. Third, the results of the comparison between the groups show that the factors that influence smart home adoption can be different according to the characteristics of users in the case of a smart home

service. In the case of a smart home service, the main users are often the elderly, patients, and women, and the results clearly show the effect of user characteristics on acceptance behavior. Therefore, acceptance studies that may be different according to user characteristics demonstrate that a research design needs to consider user characteristics.

This study also provides several useful insights for practitioners who manage the development or marketing of a smart home service. First, functional diversity should be assured to consumers. Due to the different smart home service requirements determined by the customer group, it is necessary to systematically support each consumer group in selecting the desired functions. For example, it is possible to provide a smart home service with high interconnectivity with other devices for apartment residents, while focusing on delivering an automation function for general home residents. To this end, smart home service companies should consider how to configure smart home services, including cooperating with third party device manufacturers for each type of customer, and prepare various plans according to the detailed function configuration. Second, continuous R&D on AI-based automation is required. It was revealed that automation has a positive impact on the intention for continuous use of current smart home users. Therefore, it is reasonable to infer that as the basic controllability-based smart home service spreads, consumers' need for automation will increase. A large-scale investment on basic infrastructure such as data centers, cloud, or big data systems by smart home service companies will be required in order to analyze customers' lifestyles and interact with their movements of emotional change. Third, to increase customer reliability, smart home service providers should adopt high-level security technologies and set up internal policies to prevent information leakage. Trust in smart home service providers has become a significant issue as data-based smart home companies are rapidly expanding, such as Google.

7. Conclusion and Further Study

This study empirically examined important factors for the adoption and spread of smart home services. Research results show that interconnectivity and reliability are required along with the right level of automation. Furthermore, because there are differences in preference factors according to the characteristics of users, it has been confirmed that the service design that considers these characteristics is necessary. If these factors are taken into consideration, smart home services that have not been activated in the past will spread and the market will grow.

Although the findings of this study provide meaningful insights into the adoption of smart home services, this study has limitations that future research should address. First, key findings of this research are based on the data only from South Korea. A future study should attempt to gather ethnically and geographically diverse data to ensure the generalizability of the results. Second, hedonic-related variables may be considered as influence factors of adoption intention such as perceived design (e.g., visual attractiveness of control hub hardware and software user interface). Lastly, in future studies, a significant difference in the antecedents' influence on behavioral intention between current and potential users may be found. Despite the limitations, this study contributes to a more systematic understanding of smart home service adoption. In this regard, we hope that this study helps to create a foundation for related future research.

Conflicts of Interest

Authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This study was supported by the BK21+Intellectual Property·Information Protection Law Expert Program.

References

[1] Y. Nakamura, Y. Arakawa, T. Kanehira, M. Fujiwara, and K. Yasumoto, "SenStick: comprehensive sensing platform with an ultra tiny all-in-one sensor board for IoT research," *Journal of Sensors*, vol. 2017, Article ID 6308302, 16 pages, 2017.

[2] Mordorintelligence, "Global smart homes market—growth, analysis, forecast to 2022," https://www.mordorintelligence.com/industry-reports/global-smart-homes-market-industry?gclid=CIjd6MXjydYCFYYDKgod4ZQFaw.

[3] V. Ricquebourg, D. Menga, D. Durand, B. Marhic, L. Delahoche, and C. Loge, "The smart home concept: our immediate future," in *2006 1ST IEEE International Conference on E-Learning in Industrial Electronics*, pp. 23–28, Hammamet, Tunisia, 2006, IEEE.

[4] W. K. Edwards and R. E. Grinter, "At home with ubiquitous computing: seven challenges," in *Ubicomp 2001: Ubiquitous Computing*, pp. 256–272, Springer, Berlin, Heidelberg, 2001.

[5] N. Balta-Ozkan, R. Davidson, M. Bicket, and L. Whitmarsh, "Social barriers to the adoption of smart homes," *Energy Policy*, vol. 63, pp. 363–374, 2013.

[6] R. Lutolf, "Smart home concept and the integration of energy meters into a home based system," in *Seventh International Conference on Metering Apparatus and Tariffs for Electricity Supply*, pp. 277-278, Glasgow, UK, 1992, IET.

[7] C. Kidd, R. Orr, G. Abowd et al., "The aware home: a living laboratory for ubiquitous computing research," in *Cooperative Buildings. Integrating Information, Organizations, and Architecture*, pp. 191–198, Springer, Berlin, Heidelberg, 1999.

[8] A. M. Adami, T. L. Hayes, and M. Pavel, "Unobtrusive monitoring of sleep patterns," in *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No.03CH37439)*, vol. 2, pp. 1360–1363, Cancun, Mexico, 2003, IEEE.

[9] H. Andoh, K. Watanabe, T. Nakamura, and I. Takasu, "Network health monitoring system in the sleep," in *SICE 2004 Annual Conference*, vol. 2, pp. 1421–1424, Sapporo, Japan, 2004, IEEE.

[10] T. Koskela and K. Väänänen-Vainio-Mattila, "Evolution towards smart home environments: empirical evaluation of three user interfaces," *Personal and Ubiquitous Computing*, vol. 8, no. 3-4, pp. 234–240, 2004.

[11] A.-G. Paetz, B. Becker, W. Fichtner, and H. Schmeck, "Shifting electricity demand with smart home technologies—an

experimental study on user acceptance," in *30th USAEE/IAEE North American Conference*, vol. 19, p. 20, Washington, DC, USA, 2011.

[12] A. Leeraphong, B. Papasratorn, and V. Chongsuphajaisiddhi, "A study on factors influencing elderly intention to use smart home in Thailand: a pilot study," in *The 10th International Conference on e-Business*, Bangkok, Thailand, 2015.

[13] A. Alaiad and L. Zhou, "Patients' behavioral intentions toward using WSN based smart home healthcare systems: an empirical investigation," in *2015 48th Hawaii International Conference on System Sciences*, pp. 824–833, Kauai, HI, USA, 2015, IEEE.

[14] X. Fan, Q. Xie, X. Li et al., "Activity recognition as a service for smart home: ambient assisted living application via sensing home," in *2017 IEEE International Conference on AI & Mobile Services (AIMS)*, pp. 54–61, Honolulu, HI, USA, 2017, IEEE.

[15] L. Vadillo, M. L. Martín-Ruiz, I. Pau, R. Conde, and M. Á. Valero, "A smart telecare system at digital home: perceived usefulness, satisfaction, and expectations for healthcare professionals," *Journal of Sensors*, vol. 2017, Article ID 8972350, 12 pages, 2017.

[16] H. Bao, A. Y. L. Chong, K. B. Ooi, and B. Lin, "Are Chinese consumers ready to adopt mobile smart home? An empirical analysis," *International Journal of Mobile Communications*, vol. 12, no. 5, pp. 496–511, 2014.

[17] E. Park, Y. Cho, J. Han, and S. J. Kwon, "Comprehensive approaches to user acceptance of Internet of Things in a smart home environment," *IEEE Internet of Things Journal*, vol. 4, no. 6, pp. 2342–2350, 2017.

[18] J. Ji, T. Liu, C. Shen et al., "A human-centered smart home system with wearable-sensor behavior analysis," in *2016 IEEE International Conference on Automation Science and Engineering (CASE)*, pp. 1112–1117, Fort Worth, TX, USA, 2016, IEEE.

[19] S. Tanwar, P. Patel, K. Patel, S. Tyagi, N. Kumar, and M. Obaidat, "An advanced Internet of Thing based security alert system for smart home," in *2017 International Conference on Computer, Information and Telecommunication Systems (CITS)*, pp. 25–29, Dalian, China, 2017, IEEE.

[20] E. Elmasllari and A. Al-Akkad, "Smart energy systems in private households: behaviors, needs, expectations, and concerns," in *2017 IEEE 14th International Conference on Networking, Sensing and Control (ICNSC)*, pp. 152–157, Calabria, Italy, 2017, IEEE.

[21] A. La Marra, F. Martinelli, P. Mori, and A. Saracino, "Implementing usage control in internet of things: a smart home use case," in *2017 IEEE Trustcom/BigDataSE/ICESS*, pp. 1056–1063, Sydney, NSW, Australia, 2017, IEEE.

[22] H. Ghayvat, S. Mukhopadhyay, X. Gui, and N. Suryadevara, "WSN- and IOT-based smart homes and their extension to smart buildings," *Sensors*, vol. 15, no. 5, pp. 10350–10379, 2015.

[23] S. Sintonen and M. Immonen, "Telecare services for aging people: assessment of critical factors influencing the adoption intention," *Computers in Human Behavior*, vol. 29, no. 4, pp. 1307–1317, 2013.

[24] R. Parasuraman and V. Riley, "Humans and automation: use, misuse, disuse, abuse," *Human Factors*, vol. 39, no. 2, pp. 230–253, 1997.

[25] C. Reinisch, M. J. Kofler, and W. Kastner, "ThinkHome: a smart home as digital ecosystem," in *4th IEEE International Conference on Digital Ecosystems and Technologies*, pp. 256–261, Dubai, UAE, 2010, IEEE.

[26] A. Geraci, F. Katki, L. McMonegal et al., *IEEE Standard Computer Dictionary: Compilation of IEEE Standard Computer Glossaries*, IEEE Press, Piscataway, NJ, USA, 1991.

[27] M. Chan, D. Estève, C. Escriba, and E. Campo, "A review of smart homes—present state and future challenges," *Computer Methods and Programs in Biomedicine*, vol. 91, no. 1, pp. 55–81, 2008.

[28] M. Sági, D. Mijic, D. Milinkov, and B. Bogovac, "Smart home automation," in *2012 20th Telecommunications Forum (TELFOR)*, pp. 1512–1515, Belgrade, Serbia, 2012, IEEE.

[29] T. T. Luor, H.-P. Lu, H. Yu, and Y. Lu, "Exploring the critical quality attributes and models of smart homes," *Maturitas*, vol. 82, no. 4, pp. 377–386, 2015.

[30] R. Kalman, "On the general theory of control systems," *IRE Transactions on Automatic Control*, vol. 4, no. 3, p. 110, 1959.

[31] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, "Controllability of complex networks," *Nature*, vol. 473, no. 7346, pp. 167–173, 2011.

[32] J. Choi, H. J. Lee, F. Sajjad, and H. Lee, "The influence of national culture on the attitude towards mobile recommender systems," *Technological Forecasting and Social Change*, vol. 86, no. 1, pp. 65–79, 2014.

[33] K. Lyytinen and Y. Yoo, "Research commentary: the next wave of nomadic computing," *Information Systems Research*, vol. 13, no. 4, pp. 377–388, 2002.

[34] C. Kühnel, T. Westermann, F. Hemmert, S. Kratz, A. Müller, and S. Möller, "I'm home: defining and evaluating a gesture set for smart-home control," *International Journal of Human-Computer Studies*, vol. 69, no. 11, pp. 693–704, 2011.

[35] C. Roduner, M. Langheinrich, C. Floerkemeier, and B. Schwarzentrub, "Operating appliances with mobile phones—strengths and limits of a universal interaction device," in *International Conference on Pervasive Computing*, pp. 198–215, Toronto, ON, Canada, 2007, Springer.

[36] M. Pagani, "Determinants of adoption of third generation mobile multimedia services," *Journal of Interactive Marketing*, vol. 18, no. 3, pp. 46–59, 2004.

[37] H. Yang, J. Yu, H. Zo, and M. Choi, "User acceptance of wearable devices: an extended perspective of perceived value," *Telematics and Informatics*, vol. 33, no. 2, pp. 256–269, 2016.

[38] A. H. Maslow, "A theory of human motivation," *Psychological Review*, vol. 50, no. 4, pp. 370–396, 1943.

[39] P. Keen, G. Ballance, S. Chan, and S. Schrump, *Electronic Commerce Relationships: Trust by Design*, Prentice Hall PTR, Upper Saddle River, NJ, USA, 1999.

[40] M. Friedewald, O. D. Costa, Y. Punie, P. Alahuhta, and S. Heinonen, "Perspectives of ambient intelligence in the home environment," *Telematics and Informatics*, vol. 22, no. 3, pp. 221–238, 2005.

[41] W. W. Chin and J. Dibbern, "An introduction to a permutation based procedure for multi-group PLS analysis: results of tests of differences on simulated data and a cross cultural analysis of the sourcing of information system services between Germany and the USA," in *Handbook of Partial Least Squares*, pp. 171–193, Springer, Berlin, Heidelberg, 2010.

[42] J. C. Augusto and C. D. Nugent, "Smart homes can be smarter," in *Designing Smart Homes*, pp. 1–15, Springer, Berlin, Heidelberg, 2006.

[43] S. Hammer, M. Wißner, and E. André, "Trust-based decision-making for smart and adaptive environments," *User Modeling and User-Adapted Interaction*, vol. 25, no. 3, pp. 267–293, 2015.

[44] M.-H. Hsu and C.-M. Chiu, "Internet self-efficacy and electronic service acceptance," *Decision Support Systems*, vol. 38, no. 3, pp. 369–381, 2004.

[45] B. Van der Vlist, G. Niezen, J. Hu, and L. Feijs, "Design semantics of connections in a smart home environment," in *Proceedings of Design and Semantics of Form and Movement (DeSForM)*, pp. 48–56, Lucerne, Switzerland, 2010.

[46] H. Yang, H. Lee, and H. Zo, "User acceptance of smart home services: an extension of the theory of planned behavior," *Industrial Management & Data Systems*, vol. 117, no. 1, pp. 68–89, 2017.

[47] P. Meso, P. Musa, and V. Mbarika, "Towards a model of consumer use of mobile information and communication technology in LDCs: the case of sub-Saharan Africa," *Information Systems Journal*, vol. 15, no. 2, pp. 119–146, 2005.

[48] F. D. Davis, R. P. Bagozzi, and P. R. Warshaw, "User acceptance of computer technology: a comparison of two theoretical models," *Management Science*, vol. 35, no. 8, pp. 982–1003, 1989.

[49] W. Bhuasiri, H. Zo, H. Lee, and A. P. Ciganek, "User acceptance of e-government services: examining an e-tax filing and payment system in Thailand," *Information Technology for Development*, vol. 22, no. 4, pp. 672–695, 2016.

[50] J. C. Anderson and D. W. Gerbing, "Structural equation modeling in practice: a review and recommended two-step approach," *Psychological Bulletin*, vol. 103, no. 3, pp. 411–423, 1988.

[51] C. Fornell and D. F. Larcker, "Structural equation models with unobservable variables and measurement error: algebra and statistics," *Journal of Marketing Research*, vol. 18, no. 3, pp. 382–388, 1981.

Hindawi
Journal of Sensors
Volume 2018, Article ID 3202787, 15 pages
https://doi.org/10.1155/2018/3202787

Research Article

Real-Time Cloud-Based Health Tracking and Monitoring System in Designed Boundary for Cardiology Patients

Aamir Shahzad[1], **Yang Sun Lee**[2], **Malrey Lee**[3], **Young-Gab Kim**[1], **and Naixue Xiong**[4]

[1]*Department of Computer and Information Security, Sejong University, 209 Neungdong-ro, Gwangjin-gu, Seoul 05006, Republic of Korea*
[2]*Division of Convergence Computer & Media, Mokwon University, Deajeon, Republic of Korea*
[3]*Advanced Image and Information Technology, School of Electronics & Information Engineering, Chonbuk National University, 664-14 1Ga, Deokjin-Dong, Jeonju 561-756, Republic of Korea*
[4]*Department of Mathematics and Computer Science, Northeastern State University, 611 N. Grand Ave No. 323, Webb Center, Tahlequah, OK 74464, USA*

Correspondence should be addressed to Young-Gab Kim; alwaysgabi@sejong.ac.kr and Naixue Xiong; xiongnaixue@gmail.com

Received 24 November 2017; Accepted 28 February 2018; Published 21 May 2018

Academic Editor: Ka L. Man

Telemonitoring is not a new term, in information technology (IT), which has been employed to remotely monitor the health of patients that are located not in common places, such hospitals or medical centers. For that, wearable medical sensors, such as electrocardiography sensors, blood pressure sensors, and glucometer, have commonly been used to make possible to acquire the real-time information from the remotely located patients; therefore, the medical information is further carried, via the Internet, to perform medical diagnosis and the corresponding treatments. Like in other IT sectors, there has been tremendous progress accounted in medical sectors (and in telemonitoring systems) that changes the human life protection against several chronic diseases, and the patient's medical information can be accessed wirelessly via Wi-Fi and cellular systems. Further, with the advents of cloud computing technology, medical systems are now more efficient and scalable in processing, such as storage and access, the medical information with minimal development costs. This study is also a piece of enhancement made to track and monitor the real-time medical information, bounded in authorized area, through the modeling of private cloud computing. The private cloud-based environment is designed, for patient health monitoring called bounded telemonitoring system, to acquire the real-time medical information of patients that resided in the boundary, inside medical wards and outside medical wards, of the medical center. A new wireless sensor network scenario is designed and modeled to keep or monitor the patients' health information whole day, 24 hours. This research is a new secured sight towards medical information access and gives directions for future developments in the medical systems.

1. Introduction

In medical sectors, employing of pervasive medical devices and their connectivity with the advance networks or/and the Internet brought the new visions for human medical diagnoses, treatments and monitoring, wireless body area network (WBAN), and remote monitoring of patients' health. The pervasive devices or medical sensors are connected with the specific parts of patients' bodies, to measure the acquired medical information such as blood pressure, sugar level, heart rates, and other medical signals, and the observed medical information will be transmitted to the medical assistance or medical advisor, through the connectivity of wireless media including cellular networks, where the received medical information will be examined for further diagnosis. Automated medical analytic tools, such as electrocardiogram analyzers, are also available for medical information analyses in real-time manners and are accounted as a part of the telemonitoring system [1]. As telemonitoring systems are not the new technological solutions in the monitoring of patients' health, several medical healthcare systems [2] have been deployed to monitor the indoor

or/and remotely located patients' health status to overcome the emergency cases and to fight against and diagnose the critical diseases before they become worst. To be more enhanced, the technologies called cloud computing systems are efficient and scalable solutions for existing networks and have been playing tremendous roles in healthcare systems in terms of information monitoring, acquisition, and storage. Through employing and deploying of public cloud computing technology for healthcare, the overall processing of healthcare systems is much efficient and easy to manage. It means that the hospital can just use the services of public cloud to underpin the continuum healthcare and also can manage the administration and other required IT requirements that have the potential to retrieve the real-time information of patients without any delay, truly synchronized and securely shared among the systems (and users), and scalable in cases of workload, and the information is always be accessible when required. In inclusion, through employing of cloud computing infrastructure, a health organization is able to manage its overall organizational structure, as a solution emerges with the coherent-optimal healthcare system. Further, in cloud computing, the public cloud is efficient in the monitoring of information and managing of services, for pervasive healthcare systems and be deemed appropriate for the issue of scalability and security. As the retrieving of medical information has great value and requires to be secured during transmission over the Internet, therefore several medical organizations have not been attaining upon the public cloud platforms, due to the security issues and to gain the high-security level during information exchanges. In short, public cloud services are efficient and reliable, but they still encompass potential vulnerabilities because they are called publicly open for all. To resolve the issues of security, the private cloud computing is a significant and trusted solution for medical information exchanges, with the satisfactions of information confidentiality and authorized access; moreover, medical organizations can also leverage their other important resources [2–5]. Figure 1 illustrates the typical structure of cloud computing for healthcare systems.

Eucalyptus, called "Elastic Utility Computing Architecture for Linking Your Programs to Useful Systems," mainly deploys for private and hybrid cloud computing services and is an open software used to implement infrastructure as a service (IaaS), network as a service, and storage as a service, and these services are made available for end-users by consolidating with other existing infrastructures in cloud computing environment [6, 7]. In [8], ECG analyses were performed by hosting application (or analysis software), as software-as-a-service (SaaS) and as a web service, in public cloud computing environment. In cloud computing environment, the platform-as-a-service (PaaS) layer used three main modules, such as "container scaling manager or tomcat container, workflow engine as service platform, and Aneka application platform," to manage the processing of hosted software and to fulfill the demands/requests from users. Therefore, upon user requests, container scaling manager scales the number of containers to process the user queries toward workflow engine that controls the overall processing and jobs assigned to the medical software workflow and further processes them to PaaS middleware called Aneka. Aneka is an innovated technology based on Microsoft.NET platform, an application development platform (for distributive applications) which provides run-time development scenarios and application programmable interfaces to develop customized applications, on either private or public cloud computing platform [9]. Aneka software, a development platform, provides several services in the development phases of distributive applications, for example, defines useful logic for distributive applications. Moreover, it is a generic infrastructure that manages the processing of distributive applications and is a consolidation platform to work with other existing infrastructures (e.g., "Amazon EC2" as part of cloud computing [10]. Each workflow engine was hosted, as a part, inside a tomcat container and would be increased in numbers based on the work demands from the users. Sooner, the user jobs are being collected by the workflow engine and processed to the Aneka, where each job resided in the waiting queue is triggered depending on the availability of resources. Therefore, by using of the master-worker scenario, Aneka manages the upcoming communication between the IaaS that runs Aneka master service and Aneka workers, installed in VM at PaaS. Furthermore, waiting-request scheduling at the workflow and the average user-requests handled by the container in specific session are performed in run-time through cloud load balancing [8, 9].

As concluded, the most efficient way to manage an organization's dedicated resources, especially from the encounter of security, and keep their information internally as part of organizational regulations without enrolments of external entities (e.g., outside of the organization), is private computing platform. Cloud computing provides and manages an elastic pool of resources based upon requirements by the end-users' perspectives; private cloud computing employs in case the resources that are designated for a single organization and public cloud computing in the case of multiply organizations shared the resource pool. The difference between the public and private computing is very little; for example, large scaling of Internet services and applications are mainly managed through employing of public cloud and the virtualization technology is utilized through private cloud computing; but at the same time, they have some very common characteristics, such as pool of resources, independent access, resource elasticity and management, and resource metering [6].

For the life demands and to improve and to fulfill the medical requirements of human lives, the notable changes have been accounted in medical sectors that are significant not only for the indoor patients but also for the remote-home residential patients. The use of wireless sensor devices and sensor networks made this possible to provide the medical cares and monitor the facilities for the remote patients. The continuous monitoring of remote (located) patients' status (i.e., blood pressure, heart rates, sugar level, and others) is significant in order to detect the critical diseases before the emergency conditions happening and also provides various medical services for the caregivers to manage the risk cases of patients and their cognitive disorders, while happening

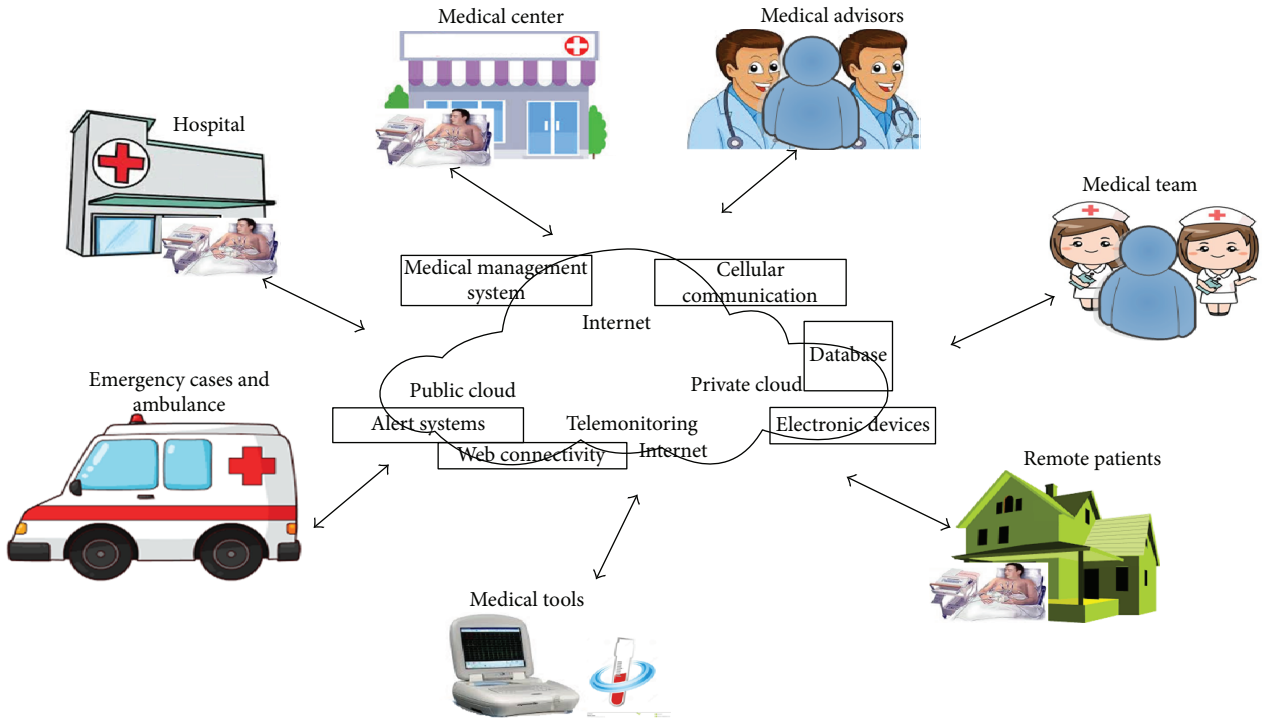


Figure 1: Typical cloud computing structure for healthcare systems.

[11, 12]. Usually, the medical sensor devices are embedded or placed on specific parts of the patient where the medical rates (or information) are required to measure, and the observed information is further transmitted wirelessly, monitored, and stored at the designated controller site (or computer machine) where the authorized persons (or authorized medical advisor) are situated for information analyzing and diagnosis purposes [12]. In last decades, a number of researches have been conducted for the development of advance context-aware medical applications, by a combination of computer networks, cellular networks (e.g., 2G, GRPS, and 3G), and wireless sensor networks. On the other side, with the great development of unobtrusive sensor devices, Bluetooth devices, and radio frequency identification- (RFID-) based sensor devices, for medical systems, their automated configurations with wireless networks make it possible to process the patients' acquisition information, from anywhere, from anyplace. These enhancements made are accounted as tremendous changes in medical systems and in reducing of critical sudden conditions of patients during an emergency and in reducing the effects of diseases, as the diseases are diagnosed and handled before going to be critical. The remote monitoring of patients' health or telemonitoring systems via sensor devices and the use of wireless sensor networks play an important role in making a coordination between the patients and medical advisor and also between medical sensor devices and other parts of the telemonitoring system [11–15].

Upon the requirements, which commonly specified by health organizations and the health center (or hospital) in our case, building a private healthcare system also mainly targets the security obstacles as its main goal. The idea to deploy a private cloud platform, for the healthcare system, is significant in ways to reduce the cost and usage of the scalable network and the available resources [6, 8, 16]. In this study, through the proposed solution to access and monitor the medical information from patients in cases of the inward system and outward system, private cloud computing platform is designed by employing of virtualization technology in Microsoft.Net platform which allows to design and employ the private cloud computing platform, for medical information monitoring, analyzing; and automated classification, as infrastructure-as-a-service. In fact, the main goal of this study is to access and monitor the real-time information of cardio-patients through the usage of built-in ECG belt having embedded sensor and ECG device where required. Therefore, the private cloud platform is modeled only that would easy to make integration with other modules of the proposed system such as cellular module, ECG acquisition module, Web application, and security and secure logon.

The rest of this research paper is organized as follows. Detailed literature is conducted in Section 2, and Section 3 describes the proposed system design and modeling, including private cloud design, inward monitoring, and outward monitoring. Study results are conducted in Section 3, and discussions are made. At the end, in Section 4, overall work study is concluded and future work is described.

2. Literature Survey

As the enhancements are made in telecommunication sectors, the telemedicine systems are discovered and as the passage of time, these systems have also been updated with the new technologies (i.e., telemonitoring systems) which are

significant in case of disease diagnosis and patient medication and remote monitoring of patients and treatments; by this way, the patients can be monitored from the places (e.g., medical centers and homes) where they are situated, over the cellular platform. For that, depending upon the demands, a new system is required that not only provides healthcare services inside the hospitals or healthcare centers but also provides efficient healthcare facility to remotely located patients. Therefore, the telemonitoring system is the best solution which provides medical monitoring facilities for the distance located patients and for the medical advisors. Especially, in the situation to fight against cardiology or arrhythmia that will always happen suddenly—if it happened—therefore, a possibility to monitor it is through remote networked portable ECG telemonitoring system [17–20]. In case of emergency, ECG-measured information is observed and simultaneously transmitted through cellular transmission, satellite transmission, and other computer wired/wireless networks to the control center or server, where the observed information will further be assessed for diagnosis and storage purposes [19, 20]. Usually, in telemonitoring systems, the carried information is saved into two storage modes: real-time mode, which stored information in real-time manners as its monitoring from ECG devices and in store-and-forward mode, where the carried (monitoring) information is latterly stored, after the specific session, and used for health analysis purposes [21].

The sensor devices and their utilization have prominent placed, in order of patient's medical treatments, in hospitals or/and healthcare centers. The sensing technology, employed in hospitals, is not very old (i.e., ECG or EKG) but has been changed as the passage of time corresponding to the technology enhancements and medical system requirements. Nowadays, various medical sensors have been used in hospitals in order to keep the physiological and physical status of patients, patient health information is tracked in real-time manners and diagnosis, and corresponding treatments are applied. The transducer devices are mainly used with medical sensors that detect the electric, genetic, and other signals, with the initialization of physiological signals combined with signal processing algorithms that predict the future status or behaviors of patient's health. Besides this, several meaningful applications, such as "monitoring in mass-casualty disasters, vital sign monitoring in hospitals, at-home and mobile aging, assistance with motor and sensory decline, and large-scale on-field medical and behavioral studies," relevant with healthcare systems were reviewed and major security challenges were also deemed that might be existing in wireless sensor networks (WSNs) to ensure the future privacy of medical sensitive information [22–24].

In a study [19], a new RFID tag called "miTag (medical information tag)" was introduced and employed during emergency cases or health disaster situations. This was a cost-effective solution, by employing tag, that provided inclusive information by tracking and monitoring the medical conditions of patients in each step (or state) of disaster, such as from the place where the disaster happened to the ambulance and from the ambulance to the hospital. Whereas, several controllers were designated and were able to view the overall information while connected to wide area network (WAN), to manage the number of patients in emergency units. The design of miTag was deemed highly scalable and extensive that can be used with a variant of communication systems, that is, GPS, and able to sense the patient status such as blood pressure rate, body temperature, and ECG signals. The miTag device was used and supported to transmit information in two-way directions. In the disaster scenarios, several major requirements, such as scalable usage of sensors and networks, reliable and efficient transmission ways, and costless and efficient hardware devices, have to be considered in order to track the remotely located patients' information. For that, few commercial devices are available, but most of them have a number of limitations, such as network setup and configuration issues; tracking is only limited to persons unable to identify the disaster locations, a poor installation, and integration with other systems, and so on, which therefore significantly affects the overall transmission [12, 16, 25, 26].

2.1. Case Study: ECG Telemonitoring System via Cellular Platform. In Figure 2, a telemonitoring system is proposed, which monitors the cardiology patient's information through connecting with ECG (electrocardiogram) device, in real-time mode and store-and-forward mode which encompassed a Holter and a controller (or server). While required, patient information is delivered via MMS (multimedia messaging service) through GPRS (general packet radio service)/GSM (global system for mobile communication) networks. Holter is a patient-worn unit where patients are situated, and ECG devices are connected with them and provided GPS (global positioning system) information which would be helpful in cases of emergency. Holter is also responsible for patients' ECG information, its storage, and assessment. In the real-time mode, patients' abnormal heartbeats rates are measured, to make be efficient, in order to measure the accurate heart rates. To do this, an algorithm is deployed that made the classifications, on measurements, of heart rates which would be significant in finding the abnormal heart rates, and based on the measured rates, the corresponding medical advisor would be concerned through MMS, in real-time manners. MMS is a new technology and developed under the 3GPP (third generation partnership project), which made it possible to transmit the multimedia messages (e.g., text, audio, and video) that varied in sizes. As the information is measured from ECG device which is directly attached to the patient, and the measured information further assessed (or classified), with the classification algorithm, resulted in the abnormal heart rates, then at the same time, this information is transmitted to the server in the form of MMS, with the embedded information of Holter-GPS. GSM/GPRS networks are utilized for MMS, and further, the patient's monitored information is transmitted to the MMS center to the server via TCP/IPs. Upon received information, server stored this patient (monitored) information and autostreamed this information to the medical advisors. Authorized medical advisor analyzes the received information, generates the corresponding feedback, and is authorized to send MMS back to the server to keep connected with the patient for the prospective information. In server end, GIS (geographic information system) software is installed that is used to locate the patients

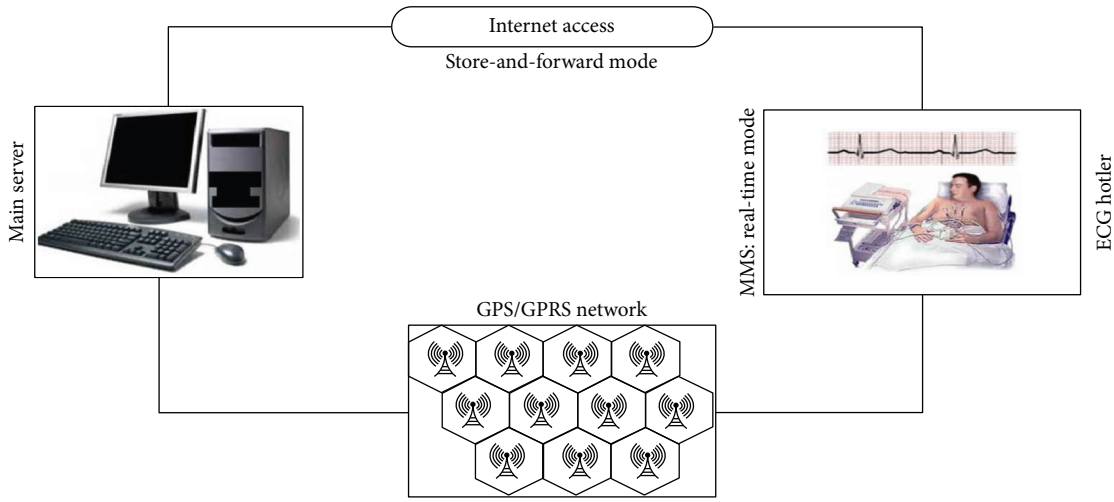


Figure 2: Telemonitoring system via cellular access [21].

in emergency cases; as the delivered MMS was packed with GPS information, therefore, this information is further used to locate the patient's location (i.e., Holter location). At both sides, in either MMS sent from the patient or sent from the medical advisor, the communication always carried in real-time senses. Real-time mode is more reliable, as the MMS (e.g., patient abnormal heartbeat rates) is transmitted, stored, and analyzed in minimal time, or without any time delay. While in case of store-and-forward mode, the Holter measured and stored the patient ECG data continuously up to 48 hours and sent the monitored information to the server for storage and future analysis purposes. In this case, information is carried through the internet, in the presence of wired networks and TCP/IPs [21].

As analyzed from above, considerable changes have been accounted in medical systems due to the gained and deployed wireless sensor technologies. The existing works [18, 19] were also conducted on to monitor the patient's status or conditions by means of various detection approaches such as fall detection, gesture and posture detections, and GPS-based tracking for patients and through several medical sensors and RFID tags that have been very commonly employed in telemonitoring systems [21]. The telemonitoring systems have been playing very important roles in reducing the effects of sudden critical diseases (i.e., arrhythmia causes) and give future trends for pervasive healthcare systems. With those systems, several benefits will be achieved and with these benefits, the patients should able to track and monitor the risky situations during emergency. Furthermore, the children and old people, who might have to suffer from cognitive and chronic illness (and other physical disorders), can also be attained in order to reduce the effects of their diseases, with safer, happier, and more independent lives [19, 21, 24]. To make a new extension as a part in healthcare systems or telemonitoring systems, the technology called cloud computing is required in which cardiology patients will be undertaken and monitored through the RIFD-based medical tags that are considered to be implanted in patients' bodies (i.e., over their hearts) and real-time information, for example, heart normal and abnormal rates, will be captured, monitored, analyzed, stored, and viewed by the authorized medical advisors and also will be viewed by the family member of the patient under security authorization rules.

As a consequence, to gain the high efficiency, scalability, and security, with the overall minimal consumption of cost, most of the information technology- (IT-) based systems, such as educational, industrial, transportation, and medical systems, are now connected and communicating through cloud computing platform which minimizes the difficulties and complexities of end-user workloads, without purchasing of any expensive software or hardware. Similarly, there is also a new trend of communication, also called ambient intelligence (AMI), based on ubiquitous communication platform, a cloud computing environment that provided interaction between the end-users and various electronic or/and sensor devices over wireless sensor networks, a context-ware design which fulfill users' demands based upon their technological requirements [27].

In the proposed Cloud-IO design, three important services, telemonitoring, location identification, and communication service, were taken into account, which modeled in the way to be consolidated with each other and should be functional over the cloud computing environment.

(i) In telemonitoring service, the required information was gathered according to the situation and from users or patients, in premises of deployed wireless network, through the connectivity of various sensors, such as medical sensors and biometric sensor devices [27, 28].

(ii) A service, called location identification, was also accounted as a part of Cloud-IO, which provided the real-time identification of the cloud users and other network entities through estimating of their positions by deploying of location estimation algorithms. Those algorithms were efficient in managing of alarms and other indications acquired from cloud users and from configured system entities. Implementing the position estimation algorithms, such

as more advances with the uses of frequency difference of arrival (FDOA) and time difference of arrival (TDFA) [29], that compute more accuracy compared with existing and are adequate in computations both for indoor and outdoor location estimations [27, 30].

(iii) In communication service, the text and voice messages were allowed to transfer over the designed cloud platform, through the employment of various wireless sensor devices, to fulfill the desired, overall communication goals, together with the microcontrollers as well as routing schemes and queuing managing algorithms are acquired for implementation [27].

3. System Design and Modeling

In information technology (IT) sectors, cloud computing is an efficient, scalable, and costless way to employ and deploy the organization's resources that include software, platforms, and infrastructures. A number of organizations have been employing the cloud computing platform to fulfill their organizational demands, industrial, and other requirements. In case an organization wants to keep its resources well and secured in its data center, away from the others or from external entities, the best solution that has been provided by cloud technology is private cloud computing by using the concepts of visualization [6]. In private cloud computing, the term visualization has a real meaning to deploy, employ, and manage the required organizational resources or applications running, inside the virtual environment, in virtual machines (VMs), rather install these applications in distinct hardware, in an organization, that consume high capital and management cost.

As explained above, private cloud has a special meaning for several information technology (IT) applications and services, including telemonitoring systems. Therefore, in the proposed study, a private cloud computing platform, using Microsoft.Net platform and System Center 2012, is considered and deployed as an efficient and cheaper way to access and measure the continuous real-time medical status or information of patients that are located in premises of hospital, such as inside the medical ward and outside the medical wards or in specified boundary of a hospital deemed in our study cases. This solution, System Center 2012, has been a commonly used designed platform for visualization, supporting three hypervisors: Microsoft Hyper-V, VMware ESX/ESXi, and Citrix XenServer, and provides virtual machines on demands from users (e.g., company partners) [6], and is efficient in its services like supporting various applications (i.e., software and hardware), as well as monitoring, updating, and distributing among system entities. Figure 3 illustrates the proposed study design and its main components.

3.1. Private Cloud Computing Design and Modeling. For accessing and monitoring the medical information of heart patients residing in the hospital premises, private computing platform is built up by employing of System Center 2012. This system mainly encompasses few paradigms that provide the private cloud computing values for a new system development with the specified workloads and designated to the platform entities: (1) An administer entity is designated for building the cloud and perform updating process for applications that were deployed by IT experts. (2) A developer entity is designated for building the virtual environment by virtual machine (VM) concept and also authorized to provide approving process of the VM. (3) The cloud employing applications are designed and deployed by the IT experts, and (4) during an application running in private cloud computing, the system center is designated, correspondingly, for problem detection and fixing. In Figure 4, a proposed study about private cloud computing and design is illustrated with the employing of System Center 2012.

3.1.1. Private Cloud: Virtual Machine Manager. For private cloud computing, virtual machine manager (VMM) is a building block and the main management tool for managing the organizational data center. It provides the elementary services that are used in building and managing of clouds and makes the virtualization environment through virtual machines and concepts, in which the virtual cloud hosts, with resources (e.g., storage and cloud networking), are configured and managed through VMM, as part of private cloud computing design and its services. System Center 2012 provides services, such as server management service, console service, database or storage service, library service, command shell service, and security and scalability service, to deploy the virtual machine manager for efficient private cloud computing. In our case, the main four virtual machines are created corresponding to the medical wards, meaning that there are four medical wards considered, and each encompasses only ten patients for real-time medical information monitoring, and an API was installed for accessing and monitoring of the patients, while they are outside of their medical wards in a hospital area, which may consider as virtual machines, as part of private cloud computing. To do this, wireless sensor network has modeled in which a number of sensors are configured in hospital premises that are sufficient to access and monitor the patients' medical information, through employing of ECG belt and also their positions, in the boundary of the hospital. However, the proposed design inelasticity would be extendable with new sensors, if required.

3.1.2. Private Cloud: Application Controller. Application controller provides managing of the facilities, for private cloud platform, in order to manage the deployed applications, such as ECG information acquisition and storage, ECG classification algorithm, alarm indications, and API for ECG belt, and services for information access and monitoring, respectively. An integrated self-service web portal is also provided, which allows the users to make a request directed towards the designed private cloud, to deploy, configure, and manage the created VMs (e.g., a total of six virtual machines in our study cases) and other required cloud services.

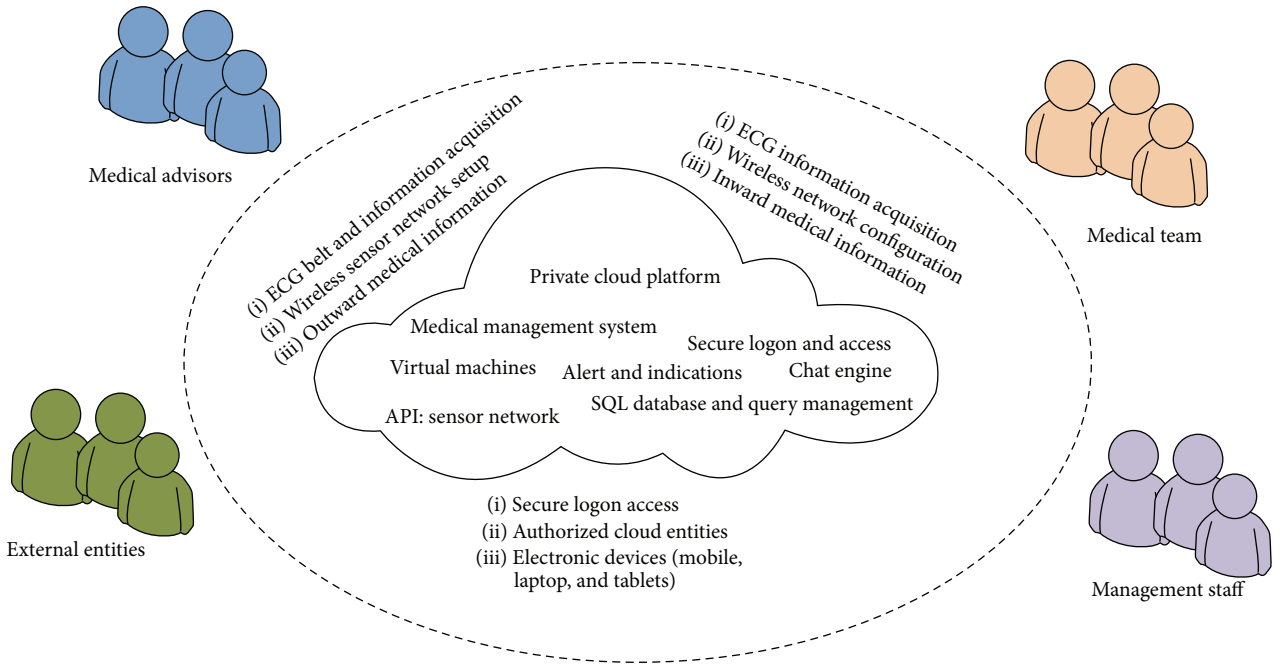


Figure 3: Proposed study design and components.

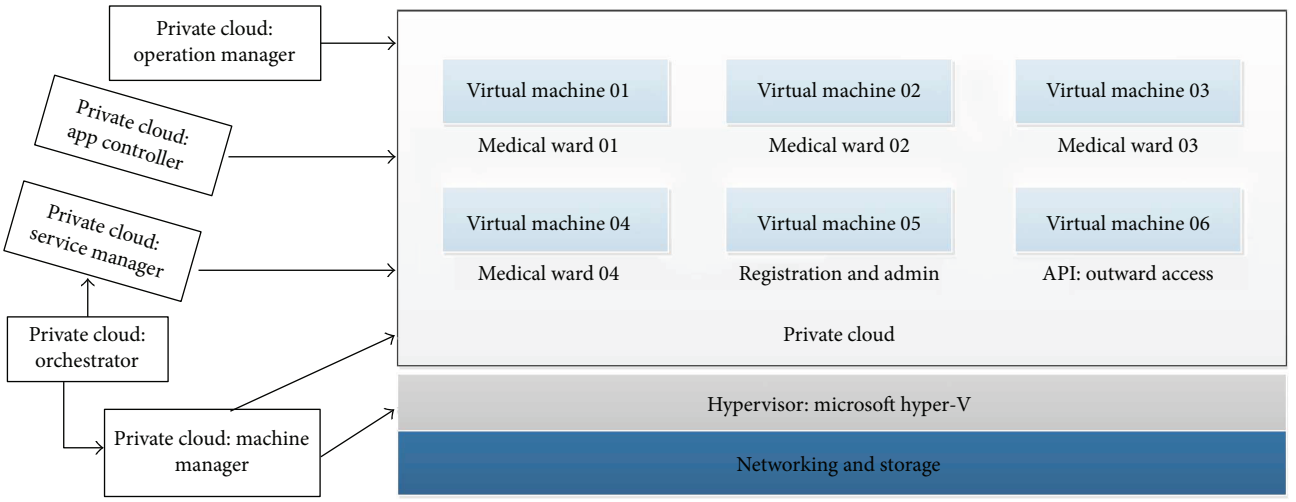


Figure 4: Proposed study: private cloud design.

3.1.3. Private Cloud: Service Manager. The service manager is an integrated, self-service portal and an automated management, as a part of private cloud computing, a platform for various services provided. For example, a new process is defined that has to manage a private computing environment, and before processing, permission criteria have to be fulfilled and should be possible with the support of service management. In our case, the external entities such as the patient's friends or family members are allowed to monitor their patients, residing in the hospital, the medical status, and other indications that will be given from the hospital management, through a secure login to the private cloud system.

3.1.4. Private Cloud: Orchestrator. In a private computing environment, an orchestrator is a solution, supported by data center, in order to manage the workflows, make interactions between employed management tools, and provide automated facilities for the users to build, implement, and monitor the resources in their own ways.

3.1.5. Private Cloud: Operation Manager. An operation manager is one of the important aspects of private cloud computing, which supports in order to monitor the running cloud virtual machines, applications, and services and perform the triggers against the issues while these happened and detected in a cloud environment.

3.2. Databases. In this study, the database is one of the important modules which keeps a record of information that are retrieved from patients through the ECG device connectivity. Database tool, called MySQL, is employed to deploy and to keep the records of each patient medical information by means of identifier that makes a distinction between each patient's captured information in his/her identity. Each time information is transmitting, in real time manners; store in the form of table's records that will further retrieve based on the queries will be requested or required by the medical individuals.

3.3. Mobile Application Design and Software. In this study, a mobile application is the most important and one of the new developments modeled to retrieve the real-time information from medical devices attached to measure the patient cardioinformation. Therefore, for this purpose, medical application is designed and modeled, which will be installed in the authorized cellular device, with the registration of a user name and security password from the cloud, which makes the automated connectivity with the external medical devices (such as ECG devices) through Bluetooth and Wi-Fi accesses. Upon logging on to the cloud, the application is installed in a cellular device and the ECG information will be viewed only in real-time manners. The other electronic devices, such as laptops and tablets, can also be used to visualize the medical information, in real-time manners, prior to register via the cloud. Among other, the main goal of this study is to get measurements from the connected ECG devices and to generate corresponding reports, within the private cloud-computing environment, which means that overall medical system access is only limited and accessible within the medical center frontier. To do this, the wireless sensor network is designed and networked within the premises of a medical center; however, the coverage that attained through signal strength covered most parts of the medical center.

In this study, permanent caregivers are the individuals, medical advisors, and medical staff, which are the main entities attached to patients' care while they are in the medical center for the short and long medical treatments. Therefore, through the cellular medical application, the detailed access is provided; the access might be limited according to the designated patients. It means that few patients are designated for one doctor and 2-3 persons from the medical staff; the cellular application provides detailed real-time information of these patients only, which would be viewed on cellular device and the detailed report viewed in the table includes overall information about the existing and current medical information of patients. As the current medical system designed is also mainly focused to employ a potential security solution against the adversaries or unauthorized individuals that might reside inside the medical center boundary or intercepting from outside network. In both cases, the security has to be managed to avoid the inside/outside attacking issues, as the medical information is always somehow critical to protect. Therefore, shorter access for few days or based on patient treatment session, authorized access will be provided to patient's caregivers, such as patient family member(s), that might be with the patient to support additional services, such buying of medicine, providing meals to the patient, and other important services which the patient needs. Therefore, caregivers or family members of a patient can access the information of his/her patient via cellular phone through a medical application installed. This is also considered significant to inform the caregiver in case of an emergency or an indication for patient services, for that they are responsible (like patients' medicine intake schedule, eating, and others, while he/she is not near to the patient, or outside the patient room or medical ward.

3.4. Chat Engine and Special Purpose Access. Chat engine facility, in Figure 5, is also a part of proposed system design which was employed to provide group chat service (or text communication) among the medical advisors appointed or working in the medical center. Through several discussions made in a group, the chat would be useful to fight against the critical diseases and to make the patient's health with more care.

The proposed system is deployed using private computing platform and is fully accessible to retrieve the overall information that only is limited to the area of medical center; however, in few critical cases, the information will also be retrievable via the cellular devices, outside of medical center boundary, or everywhere, registered under the names of medical advisors. So, using this facility, the medical advisors are allowed to view the current medical status of their designated patients continuously, even though they are also allowed to send feedback, for example, for health treatments, through the use of cellular application chat box, via the Internet. Therefore, these are significant demonstrations to avoid some major medical issues happening with the patients.

3.5. Information Capturing and Storage. In studies [31–36], developments have been made to retrieve patient's real-time medical information or medical status, over the far distance, or remotely from everywhere, which were significant in getting of information and in monitoring, as well. But in the current study based on some requirements, specifically due to the security issues, communication is limited to the local private body area networks (PHBAN), a network which retrieves the information from the patient through employing of medical noninvasive equipment or noninvasive devices attached with the patients; in our case, we use the ECG devices and also ECG belt with wireless sensor (ECGBWS) according to the scenarios, such as inward monitoring (IWM) and outward monitoring. Whereas, this study also deployed a potential security mechanism, through cryptography [37, 38], for ECGBWS protection against the unauthorized entities, and would employ the self-launched attacking scenarios that a measurement will be efficient against the insecurities and for measurement computed from security mechanism evaluation (process).

3.5.1. Inward Monitoring (IWM). Inward monitoring (IWM) has been a common development in telemonitoring systems [6, 13], which provides monitoring facility to monitor the health of those patients residing in the medical units or

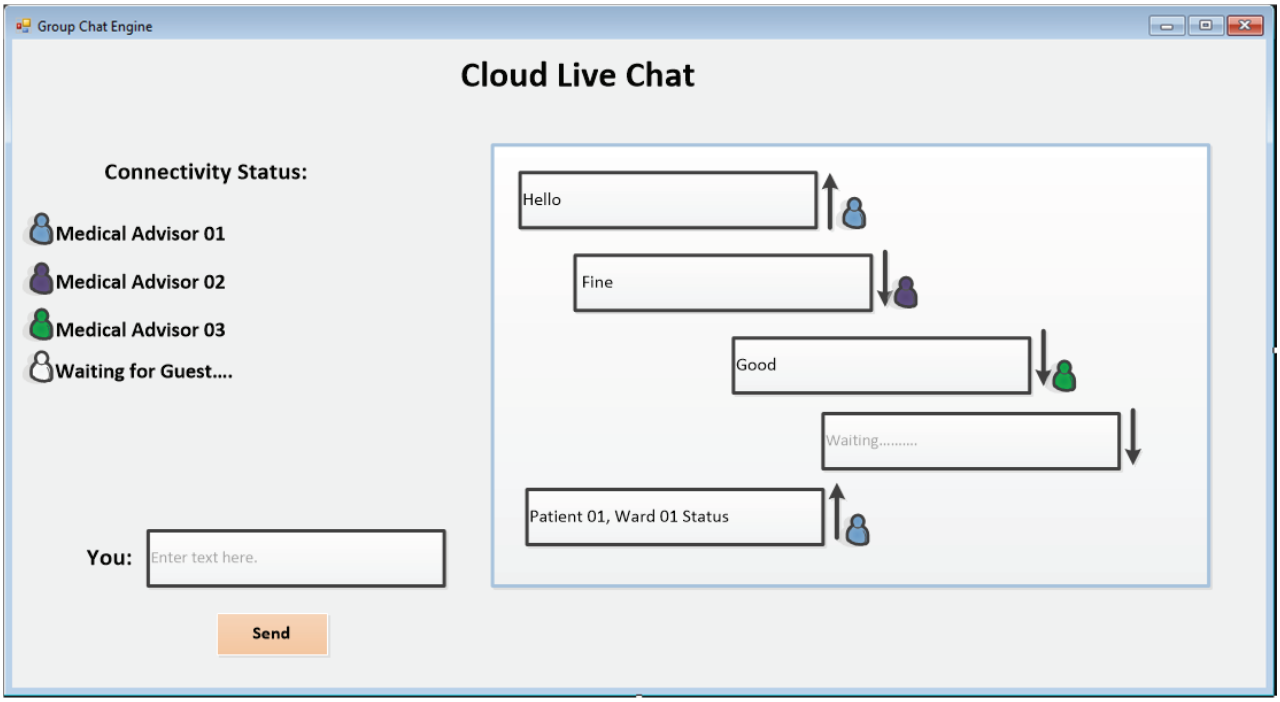


Figure 5: Cloud computing: group chat engine.

medical wards; generally, our study case is the same as those existing studies in the area of health monitoring of patient especially for arrhythmia patient. Furthermore, however, we have introduced an application design that was based on simulation, for an instance, is efficient in case the transmission from cloud computing server is disconnected due to system maintenance reasons and also in case if an unauthorized event will occur and measure the overall cloud health computing system.

In real-time monitoring, in case of inward monitoring which covers the area (e.g., width × height), the information is a continuous measuring from ECG device which attached with the selective parts of the patient body, and the 12 leads are used in our case for arrhythmia measurements. Nowadays, ECG devices are also available that have integration with wireless circuitry, designed to transmit the observed information of arrhythmia patients directly towards the system that has a configuration. For inward monitoring, in case of rhythmic patient's health information measurement and monitoring, the wired network is much better as all the patients (i.e., 10 patients) reside in one medical ward and there are only 10 main LAN cable networks to carry the real-time medical information of each patient in. Nowadays, the ECG devices are equipped with a wired and wireless connectivity; therefore, in Figure 6, to avoid the network's issues as the retrieving of patient's information is critical, the proposed study uses wired LAN connections and as well as wireless connections which are also beneficial in case of disconnection issues relevant to wired networks. Furthermore, some ECG devices are designed with built-in memory that has limited space to store medical information of patient's rhythms and should transfer to external devices through the USB port.

In one medical ward, the numbers of the patients are limited up to 10 patients, in our case of study, that much scalable to accommodate the number of patients (equal to 10); however, these numbers can be increased according to the requirements and based upon the space available. Eventually, through this scenario, several medical wards would be established based on the coming numbers of the patients in the medical center. This study database design is dynamic and would able to store and manage thousands of patient's records, in terms of their personal records and the information that will be measured from ECG devices. In detail, we entered 200 arrhythmia patients in the designed database and were also able to manage their medical information and the corresponding reports upon individual requests (by patients, medical term, and medical advisors). A deemed medical center is designed to accommodate 500 patients' based on the available space or number of built medical units.

In the medical ward, each patient resided in is accommodated with an ECG device attached to measure the continuous electrical signals or events of patients' arrhythmias, through attached leads. Overall, computing measurements, ECG waveforms that show the contractile actions of heart myocardium continuously, are observed and simultaneously displayed on embedded ECG display mainly characterized by three waves, called P, QRS, and T. In these waveforms, the QRS complex is the most important one designated to show the rhythmic activities where R is calculated and represents the peak of QRS complex. ECG is costless and an efficient noninvasive too, for diagnoses of rhythmic diseases and the continuous monitoring of the normal heart activities, in case of irregular activities or waveform intervals, changing measure in amplitude, normal time intervals, and certainly changing in heart patterns (or rhythms shapes) [2]. Further,

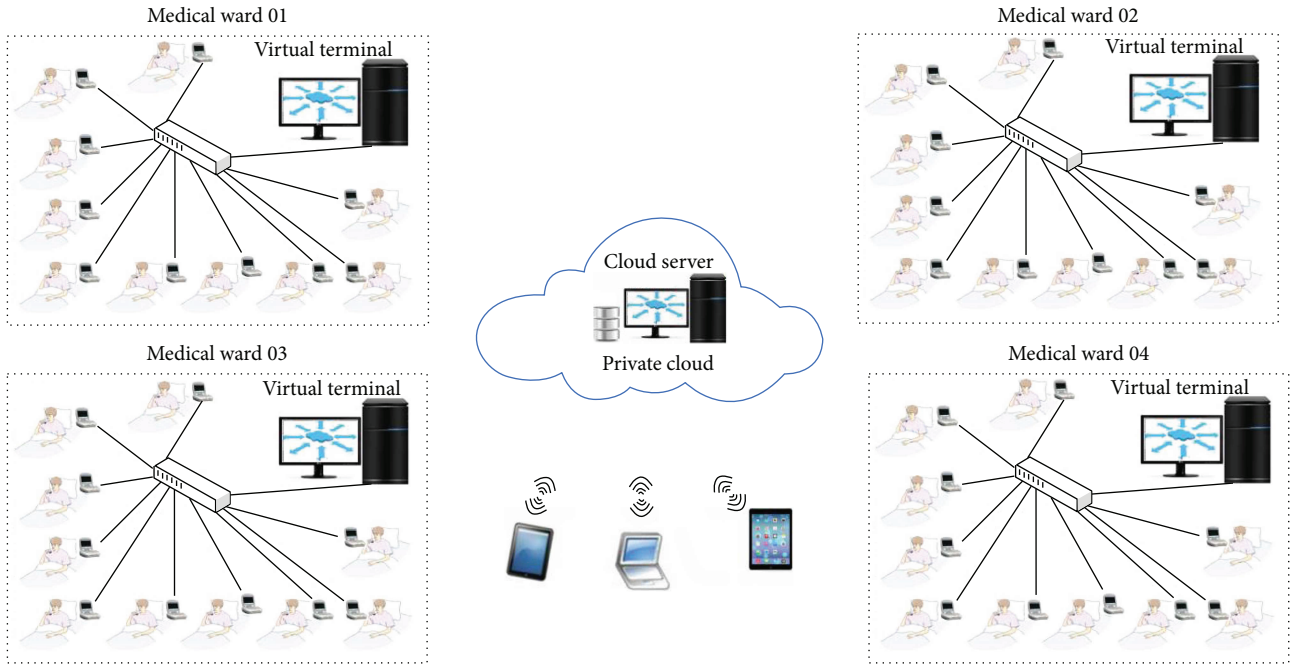


Figure 6: Inward monitoring.

the observed continuous activities of the heart, or QRS complex, are transmitted to the local unit resided in the medical ward, which means that this networked local unit is just a computer system which captures, records, and monitors the real-time rhythms of the patient illustrated in Figure 7, up to 10 patients in a medical ward, through the medical application running from the cloud server. This study is more efficient in computing of designated patients' heart health and would record this information simultaneously back to the server, without any delays of hours, days, and weeks. In Holter monitoring system, rhythms through ECG device were measured and stored in the locally located system then transmitted back after a specified session, and classification on observed medical information was performed offline, which was one of the important limitations of Holter monitoring system [21]. However, in this study, the continuous medical information of patients is captured through the ECG devices, recorded, and monitored in real-time manners, to avoid that any critical issue might happen; further analysis is performed also online through classification and analysis algorithm or application, running from cloud or cloud server. In short, the medical information will be received at the cloud server, from each patient with his/her unique identifier, the diagnostic will be performed based upon the classification algorithm implemented as part of proposed study, and simultaneously, feedback will be transmitted to the designated medical staff and also to the designated advisor for overview and for suggestions in case they will require any. The analysis was always performed based on the static computation of ECG classification algorithm installed and resided as a part of the cloud server.

(1) Application Design and Implementation. Cloud computing is an efficient, reliable, and scalable way for today's application access, more especially in telemonitoring where the medical application, such as ECG information, can retrieve or access and is hosted to monitor and analyze the retrieved medical information through analysis tools, as soft-as-a-service. Mainly, each type of application or software is hosted and can retrieve through the web-based service, where the end-users can simply request and get the desired information, such as ECG/medical information uploaded/downloaded, medical information analyses, and current and existing reports, according to their queries and demands without the knowledge of application of the underlying complexities and difficulties that might be present in the development or in hosting of any application in cloud computing platform. At a lower layer, in the top-down scenario, cloud service, platform-as-a-service, is employed to manage the processing of the application or software hosted at SaaS layer.

(2) Direct and Offline System. As explained, overall information carried from each patient connected with ECG device is transmitted to a local station located as a part of a medical ward to store and monitor the information in real-time manners through the medical application or virtual machine, running from a cloud server, which means that carried information is directly stored in the cloud server via local system and monitoring is also performed in real time to avoid the critical cases to happen. At the same time, the proposed system design is also efficient to store the medical information in case of system disconnectivity, system connection failure, or other network or system internal issues; so if these cases might happen, then the offline system would be functional. In the offline system, the medical running application is designed with efficiency to receive the upcoming medical information from each patient and store it into a local memory that configures as a part of application software. While

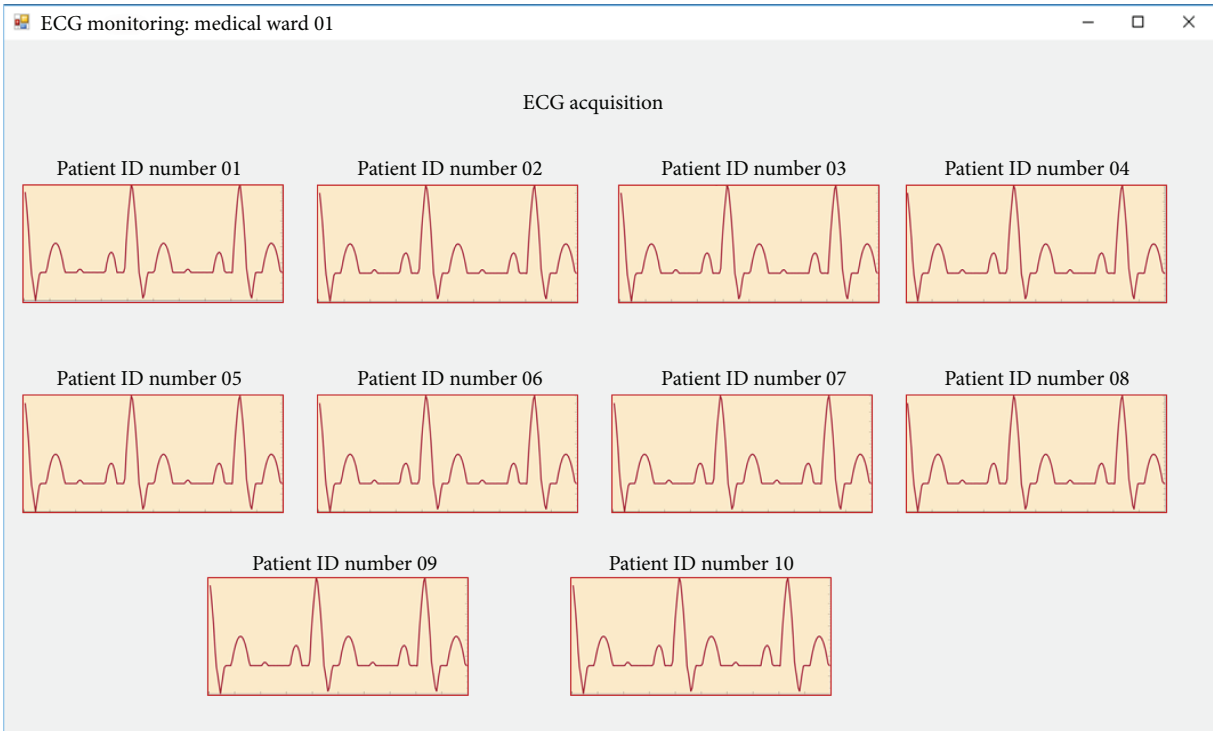


Figure 7: ECG inward monitoring.

server connection will again establish, all stored medical information will be updated in the server site database. The offline system scenario is significant in case to handle the future system failures or disconnectivity issues; also, the analysis will be performed in cases of cloud server connectivity issues. To make an analysis, a replicate copy of ECG classification algorithm is always a part of the local system, in each medical ward, that will be activated only when the cloud server will disconnect. As the server will not be in connection anymore, offline bytes or system disconnection bytes are constantly received as a part of medical application, running in the local system, within 20–40 seconds interval and the application will activate the ECG classification algorithm. As consequence, the upcoming ECG information will be stored and at the same time, the medical analysis will be performed locally in each medical ward, and the medical information will be stored and analyzed to fight against the potential critical medical events of arrhythmia.

3.5.2. Outward Monitoring (OWM). In the second phase of the proposed study, outward monitoring (OWM) is modeled to monitor the medical status of rhythm patients, while they are out from their medical wards where they most of time resided for medical treatments and also for other purposes. In most common cases, medical hardware is an equipment in different rooms within one building block or in another block, in case the medical center is surrounded by many blocks, such as blocks A, B, and C. Therefore, the movements of patients are always carried out during their medical tests, heart operations, and heart surgery and in cases of recovery corresponding with required exercises and also in cases of open-air relaxation and spending of good time of gossip with family members and friends. In short, there are many situations and medical reasons in which patients have required for moving out of their medical ward. Therefore, during that time, measuring the rhythms of the patient will be required to avoid the critical cases of arrhythmia while they are out of their medical wards. Eventually, this study proposed and modeled a sensor-based wireless network scenario to capture the ECG waveforms of patients that might be moving out of their medical ward, without connectivity with ECG devices, through the wireless ECG belt, embedded ECG sensor, fixed over the patient body.

While the patient is outside of his/her medical ward, a belt is fixed over the specific part of the patient body that has embedded ECG sensor, having a unique entity, communicating with the cloud server through the various access points (APs). A medical wireless sensor network is designed, in which sensors (or access points) are placed or networked in an area that covered most parts of the medical center. Therefore, real-time medical information could be accessed through wireless sensor configuration, and patient will be located anywhere in the boundary of the medical center or in coverage of wireless sensor network. In Figure 8, wireless sensor network is designed to carry the transmission of medical information from patients, located outside of their medical wards, through ECG sensor belt.

4. Results and Discussion

This study is based on the simulation design in which inward and outward networks were modeled to carry the medical information of arrhythmia patients in real-time fashion. As the medical information is important, therefore, it has to be

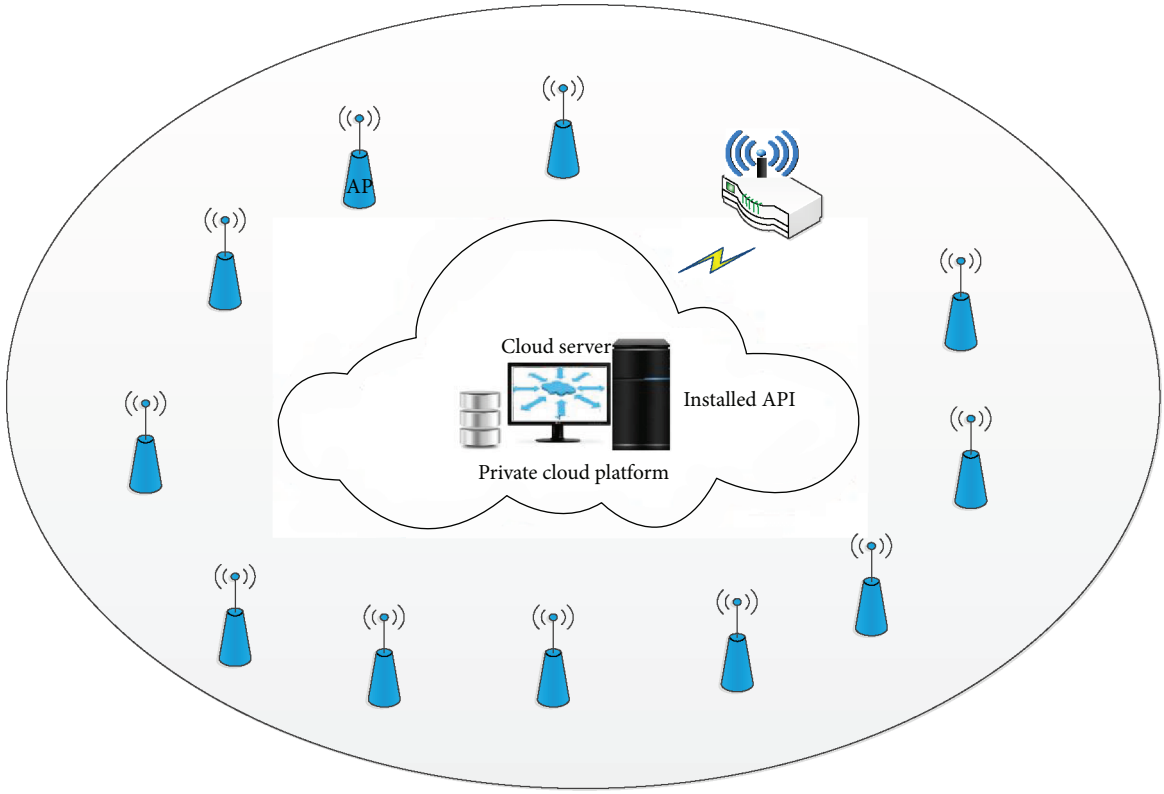


Figure 8: ECG outward monitoring.

measured continuously and in real-time manners. Overall, cloud computing designed for accessing, monitoring, and storing of arrhythmia patients' information is simulated in a desktop of a computer system, with system specifications of (1) Intel® Core™ i7-4790 CPU at a 3.60 GHz processor, (2) 8.00 GB installed memory, (3) 64-bit operating system, Windows 10, ×64-based processor, (4) Realtek PCIe GBE Family Controller, Ethernet, (5) 1000/1000 (Mbps) link speed, (6) IPv6, and others. A total of six virtual machines were deployed in private cloud system, four were installed to carry and manage the medical information coming from the medical wards (i.e., four medical wards), and the remaining virtual machines such as cloud VM number 05 and cloud VM number 06 were deployed for patient registration and administration purposes and for monitoring the medical information from ECG belt, while patients are in outward monitoring mode, through API installed in VM number 06 in private cloud.

In Table 1, the performance results are measured, in which ECG information was accessed and stored in cloud storage (i.e., in MySQL database tool), in case of inward monitoring and outward monitoring modes. A total of 50 experiments were performed to measure the results, such as ECG signals or information received and stored at the cloud server and security computations, while experiment number 00 is not mentioned in Table 1 but was performed to test the overall cloud system configuration in order to avoid the transmission errors. Each time ECG signals are received at the cloud server, the security was tested to ensure the authentication and confidentiality of that information through employing of AES cryptography algorithm [37, 38]. Therefore, the signs, such as $\surd$ and X, are used to show that the ECG information was received and the security test was verified or not. As a consequence, the performance results shown in Table 1 are significant, as high accuracy is computed based on 50 experiment results.

5. Conclusion and Future Work

As the technological demands are increasing, the medical systems have been updated with the new advanced technologies (i.e., smart medical systems); however, there are still several challenges, such as to carry the remotely located information and the remote connectivity issues, mainly present in the healthcare systems. Alongside with this, the cost is also a big issue, raising while updating the existing systems with the new developments. Therefore, the cloud computing platform is one of the best solutions in the current age to fight against the issues. In this study, based on the medical organizational specifications and requirements for communication, including the security requirements, private cloud computing environment is designed and modeled where the medical information of heart patients who underregistered or resided in the hospital, inside of the medical ward and outside of the medical ward in the premises of the hospital, is accessed and monitored in real-time manners. Further, the system's security requirements are fulfilled through the implementation of AES algorithm, which resulted in protection of overall communication in the designed private cloud computing environment. Moreover, the study provided an immense

Table 1: Performance results.

Number	Inward monitoring		Outward monitoring		Ward and patient IDs
	ECG signal	Security test	ECG signal	Security test	
1	✓	✓	✓	✓	Ward number 01, patient number 01
2	✓	✓	x	x	Ward number 01, patient number 02
3	✓	✓	✓	✓	Ward number 01, patient number 03
4	✓	✓	✓	✓	Ward number 01, patient number 04
5	✓	✓	✓	✓	Ward number 01, patient number 05
6	✓	✓	✓	✓	Ward number 01, patient number 06
7	✓	x	✓	✓	Ward number 01, patient number 07
8	✓	✓	x	x	Ward number 01, patient number 08
9	✓	✓	✓	✓	Ward number 01, patient number 09
10	✓	✓	✓	✓	Ward number 01, patient number 10
11	✓	✓	✓	✓	Ward number 02, patient number 01
12	✓	✓	✓	✓	Ward number 02, patient number 02
13	✓	✓	✓	✓	Ward number 02, patient number 03
14	✓	✓	✓	✓	Ward number 02, patient number 04
15	✓	✓	✓	✓	Ward number 02, patient number 05
16	✓	✓	✓	✓	Ward number 02, patient number 06
17	✓	✓	✓	✓	Ward number 02, patient number 07
18	✓	✓	x	x	Ward number 02, patient number 08
19	x	x	✓	✓	Ward number 02, patient number 09
20	✓	✓	✓	✓	Ward number 02, patient number 10
21	✓	✓	✓	✓	Ward number 03, patient number 01
22	✓	✓	✓	✓	Ward number 03, patient number 02
23	✓	✓	✓	✓	Ward number 03, patient number 03
24	✓	✓	✓	✓	Ward number 03, patient number 04
25	✓	✓	✓	✓	Ward number 03, patient number 05
26	✓	✓	✓	✓	Ward number 03, patient number 06
27	✓	✓	✓	✓	Ward number 03, patient number 07
28	✓	✓	✓	✓	Ward number 03, patient number 08
29	✓	✓	✓	✓	Ward number 03, patient number 09
30	✓	✓	✓	✓	Ward number 03, patient number 10
31	✓	✓	✓	✓	Ward number 04, patient number 01
32	✓	✓	✓	✓	Ward number 04, patient number 02
33	✓	✓	x	x	Ward number 04, patient number 03
34	✓	✓	✓	✓	Ward number 04, patient number 04
35	✓	x	✓	✓	Ward number 04, patient number 05
36	✓	✓	✓	✓	Ward number 04, patient number 06
37	✓	✓	✓	✓	Ward number 04, patient number 07
38	✓	✓	✓	✓	Ward number 04, patient number 08
39	✓	✓	✓	✓	Ward number 04, patient number 09
40	✓	✓	✓	✓	Ward number 04, patient number 10
41	✓	✓	✓	✓	Ward number 01, patient number 03
42	✓	✓	✓	✓	Ward number 01, patient number 06
43	✓	✓	✓	✓	Ward number 02, patient number 02
44	✓	✓	✓	✓	Ward number 03, patient number 03
45	✓	✓	✓	✓	Ward number 04, patient number 01
46	✓	✓	✓	✓	Ward number 03, patient number 01
47	✓	✓	✓	✓	Ward number 02, patient number 05

Table 1: Continued.

Number	Inward monitoring		Outward monitoring		Ward and patient IDs
	ECG signal	Security test	ECG signal	Security test	
48	✓	✓	✓	✓	Ward number 01, patient number 07
49	✓	✓	✓	✓	Ward number 03, patient number 09
50	✓	✓	✓	✓	Ward number 01, patient number 08

platform of knowledge against the causes of arrhythmia and about the arrhythmia chat engine in which medical advisors can share their ideas and suggestions that will be remarkable to fight against the enormous condemnatory heart diseases.

In early future work, this study will be extended with the new developments of hybrid cloud computing, where the main goal is to provide the medical services to the remotely located patients and the concept to measure the medical information continuously. Moreover, the study will be updated with the newly used technology in the arena of information technology called the Internet of things (IoT), which provides automated communication services over the Internet access and will be helpful for the development of new advanced medical systems.

Conflicts of Interest

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 and Communications Technology Promotion (IITP) grant funded by the Korean government (MSIT) (no. 2017-0-00756, development of interoperability and management technology of IoT system with heterogeneous ID mechanism) and the Faculty Research Fund of Sejong University in year the 2017.

References

[1] C. Doukas and I. Maglogiannis, "Bringing IoT and cloud computing towards pervasive healthcare," in *2012 Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, pp. 922–926, Palermo, Italy, July 2012.

[2] EMC Corporation, "The private cloud for healthcare enables coordinated patient care," 2010, http://middle-east.emc.com/collateral/emc-perspective/h7349-hc-private-cloud-ep.pdf.

[3] L. Heilig and S. Voss, "A scientometric analysis of cloud computing literature," *IEEE Transactions on Cloud Computing*, vol. 2, no. 3, pp. 266–278, 2014.

[4] K. Gai and S. Li, "Towards cloud computing: a literature review on cloud computing and its development trends," in *2012 Fourth International Conference on Multimedia Information Networking and Security*, vol. 146, p. 142, Nanjing, China, November 2012.

[5] P. Mell and T. Grance, *The NIST Definition of Clouding Computing Recommendations National Inst. of Standards and Technology*, vol. 145, NIST Special Publication, Gaithersburg, MD, USA, 2011.

[6] D. Chappell, "The Microsoft private cloud, a technology overview," 2011, http://www.davidchappell.com/writing/white_papers/The_Microsoft_Private_Cloud_v1.0–Chappell.pdf.

[7] M. D. Lakshmi and J. P. M. Dhas, "An open source private cloud solution for rural healthcare," in *2011 International Conference on Signal Processing, Communication, Computing and Networking Technologies*, pp. 670–674, Thuckafay, India, July 2011.

[8] S. Pandey, W. Voorsluys, S. Niu, A. Khandoker, and R. Buyya, "An autonomic cloud environment for hosting ECG data analysis services," *Future Generation Computer Systems*, vol. 28, no. 1, pp. 147–154, 2012.

[9] Y. Wei, K. Sukumar, C. Vecchiola, D. Karunamoorthy, and R. Buyya, "Aneka cloud application platform and its integration with windows azure," in *Cloud Computing: Methodology, Systems, and Applications*, L. Wang, R. Ranjan, J. Chen, and B. Benatallah, Eds., CRC Press, Boca Raton, FL, USA, 2011.

[10] C. Vecchiola, X. Chu, and R. Buyya, "Aneka: a software platform for .NET based cloud computing," *High Speed and Large Scale Scientific Computing*, vol. 18, pp. 267–295, 2009.

[11] H. Alemdar and C. Ersoy, "Wireless sensor networks for healthcare: a survey," *Computer Networks*, vol. 54, no. 15, pp. 2688–2710, 2010.

[12] A. Milenković and C. Otto, "Wireless sensor networks for personal health monitoring: issues and an implementation," *Computer Communications*, vol. 29, no. 13-14, pp. 2521–2533, 2006.

[13] R. Shahriyar, M. F. Bari, G. Kundu, S. I. Ahamed, and M. M. Akbar, "Intelligent Mobile Health Monitoring System (IMHMS)," in *Electronic Healthcare. eHealth 2009*, vol. 27 of Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, pp. 5–12, Springer, Berlin, Heidelberg, 2009.

[14] B. E. Reddy, T. V. S. Kumar, and G. Ramu, "An efficient cloud framework for health care monitoring system," in *2012 International Symposium on Cloud and Services Computing*, pp. 113–117, Mangalore, India, December 2012.

[15] M. Al-Qurishi, M. Al-Rakhami, F. Al-Qershi et al., "A framework for cloud-based healthcare services to monitor noncommunicable diseases patient," *International Journal of Distributed Sensor Networks*, vol. 11, no. 3, Article ID 985629, 2015.

[16] J. Mišić and V. Mišić, "Wireless sensor networks: performance, reliability, security, and beyond," *Computer Communications*, vol. 29, no. 13-14, pp. 2447–2449, 2006.

[17] J. G. Ko, L. Chenyang, M. B. Srivastava, J. A. Stankovic, A. Terzis, and M. Welsh, "Wireless sensor networks for healthcare," *Proceedings of the IEEE*, vol. 98, no. 11, pp. 1947–1960, 2010.

[18] R. D. Caytiles and S. Park, "A study of the design of wireless medical sensor network based u-Healthcare system,"

International Journal of Bio-Science and Bio-Technology, vol. 6, no. 3, pp. 91–96, 2014.

[19] T. Gao, C. Pesto, L. Selavo et al., "Wireless medical sensor networks in emergency response: implementation and pilot results," in *2008 IEEE Conference on Technologies for Homeland Security*, vol. 192, p. 187, Waltham, MA, USA, May 2008.

[20] N. Fernando, S. W. Loke, and W. Rahayu, "Mobile cloud computing: a survey," *Future Generation Computer Systems*, vol. 29, no. 1, pp. 84–106, 2013.

[21] C. Wen, M.-F. Yeh, K.-C. Chang, and R.-G. Lee, "Real-time ECG telemonitoring system design with mobile phone platform," *Measurement*, vol. 41, no. 4, pp. 463–470, 2008.

[22] W. Yao, C.-H. Chu, and Z. Li, "The use of RFID in healthcare: benefits and barriers," *2010 IEEE International Conference on RFID-Technology and Applications*, vol. 134, 2010, p. 128, 2010.

[23] H. Al Nahas and J. S. Deogun, "Radio frequency identification applications in smart hospitals," in *Twentieth IEEE International Symposium on Computer-Based Medical Systems (CBMS'07)*, vol. 342, p. 337, Maribor, Slovenia, June 2007.

[24] M. Mayer, N. Görtz, and J. Kaitovic, "RFID tag acquisition via compressed sensing," in *2014 IEEE RFID Technology and Applications Conference (RFID-TA)*, pp. 26–31, Tampere, Finland, September 2014.

[25] C. D. Kidd, R. Orr, G. D. Abowd et al., "The aware home: a living laboratory for ubiquitous computing research," in *Cooperative Buildings. Integrating Information, Organizations, and Architecture. CoBuild 1999*, N. A. Streitz, J. Siegel, V. Hartkopf, and S. Konomi, Eds., vol. 1670 of Lecture Notes in Computer Science, pp. 191–198, Springer, Berlin, Heidelberg, 1999.

[26] E. A. Fry and L. A. Lenert, "MASCAL: RFID tracking of patients, staff and equipment to enhance hospital response to mass casualty events," *AMIA Annual Symposium Proceedings Archive*, vol. 2005, pp. 261–265, 2005.

[27] D. I. Tapia, R. S. Alonso, Ó. García, F. de la Prieta, and B. Pérez-Lancho, "Cloud-IO: cloud computing platform for the fast deployment of services over wireless sensor networks," in *7th International Conference on Knowledge Management in Organizations: Service and Cloud Computing*, L. Uden, F. Herrera, J. Bajo Pérez, and J. Corchado Rodríguez, Eds., vol. 172 of Advances in Intelligent Systems and Computing, Springer, Berlin, Heidelberg, 2013.

[28] D. I. Tapia, R. S. Alonso, J. F. De Paz, C. Zato, and F. de la Prieta, "A telemonitoring system for healthcare using heterogeneous wireless sensor networks," *International Journal of Artificial Intelligence*, vol. 6, 2011.

[29] K. Dongkyun, J. Ha, and K. You, "Adaptive extended Kalman filter based geolocation using TDOA/FDOA," *International Journal of Control and Automation*, vol. 4, no. 2, pp. 49–58, 2010.

[30] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 37, no. 6, pp. 1067–1080, 2007.

[31] N. Amour, A. Hersi, N. Alajlan, Y. Bazi, and H. AlHichri, *Implementation of a Mobile Health System for Monitoring ECG Signals*, CiteSeerx, Princeton, NJ, USA, 2015.

[32] R. S. Tolentino, Y.-T. Kim, B. Park, and G.-C. Park, "A Design and Analysis Ubiquitous Healthcare Monitoring System over Wireless Sensor Network," *International Journal of Multimedia and Ubiquitous Engineering*, vol. 6, 2011http://www.sersc.org.

[33] D. Aranki, G. Kurillo, P. Yan, D. M. Liebovitz, and R. Bajcsy, "Real-time tele-monitoring of patients with chronic heart-failure using a smartphone: lessons learned," *IEEE Transactions on Affective Computing*, vol. 7, no. 3, pp. 206–219, 2016.

[34] D. Dziak, B. Jachimczyk, and W. J. Kulesza, "IoT-based information system for healthcare application: design methodology approach," *Applied Sciences*, vol. 7, no. 6, p. 596, 2017.

[35] V. P. Cornet and R. J. Holden, "Systematic review of smartphone-based passive sensing for health and wellbeing," *Journal of Biomedical Informatics*, vol. 77, pp. 120–132, 2018.

[36] N. Xiong, A. V. Vasilakos, L. T. Yang et al., "Comparative analysis of quality of service and memory usage for adaptive failure detectors in healthcare systems," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 4, pp. 495–509, 2009.

[37] A. Shahzad, M. Lee, H. D. Kim, S.-M. Woo, and N. Xiong, "New security development and trends to secure the SCADA sensors automated transmission during critical sessions," *Symmetry*, vol. 7, no. 4, pp. 1945–1980, 2015.

[38] S. Musa, A. Shahzad, and A. Aborujilah, "Secure security model implementation for security services and related attacks base on end-to-end, application layer and data link layer security," in *ICUIMC '13 Proceedings of the 7th International Conference on Ubiquitous Information Management and Communication*, p. 8, Kota Kinabalu, Malaysia, January 2013.

Hindawi
Journal of Sensors
Volume 2018, Article ID 1757409, 10 pages
https://doi.org/10.1155/2018/1757409

Research Article

IoT-Based Smart Building Environment Service for Occupants' Thermal Comfort

Herie Park[1] **and Sang-Bong Rhee**[2]

[1]*Automotive Lighting LED-IT Convergence Education, Yeungnam University, Gyeongsan 38541, Republic of Korea*
[2]*Department of Electrical Engineering, Yeungnam University, Gyeongsan 38541, Republic of Korea*

Correspondence should be addressed to Sang-Bong Rhee; rrsd@yu.ac.kr

Received 23 February 2018; Revised 14 April 2018; Accepted 23 April 2018; Published 21 May 2018

Academic Editor: Ka L. Man

This paper presents an Internet of Things (IoT) platform for a smart building which provides human care services for occupants. The individual health profiles of the occupants are acquired by the IoT-based smart building, which uses the accumulated knowledge of the occupants to provide better services. To ensure the thermal comfort of the occupants inside the building, we propose a dynamic thermal model of occupants. This model is based on the heat balance equation of human body and thermal characteristics of the occupants. We implement this model in two smart building models with heaters controlled by a temperature and thermal comfort index using MATLAB/Simulink®. The simulation results show that the thermal comfort-based control is more effective to maintaining occupants' thermal satisfaction and is therefore recommended for use providing human care services using IoT platforms in smart buildings.

1. Introduction

In a hyperconnected society, a human is linked to other people and machines, and machines are connected to other machines. This connectivity is formed by a physical and cybernetic linkage with the help of the sensing and communication technologies. These technologies make direct interactions among people, objects, and services that improve the productivity, efficiency, convenience, and security of our society [1]. The Internet of Things (IoT), big data, cloud computing, digital platforms, machine-to-machine (M2M) communication, artificial intelligence (AI), and machine learning are emerging technologies that support a hyperconnected society [2]. These smart technologies are applied to several domains including homes, factories, offices, transport systems, and other service and production areas. Objects and services are connected to individuals by personal communication devices and control applications that enable users to receive on-demand services.

Human care services are an example of these on-demand services. Human care services were initially explored in healthcare as a means to provide clinical services for patients who need help and care from others, such as infants and the elderly. Human care services include on-site treatment by caregivers at home or in clinical settings and long-distance treatment, also known as e-healthcare. The e-healthcare system improves the quality of patient care and reduces cost [3]. Since the e-healthcare system is composed of sensors, electronic health records, and communication protocols which are easily integrated on a commonly retailed personal device such as a smartphone, the scope of the e-healthcare services can be expanded to include the general public. For example, wearable devices such as a smartphone and a smart band have biosensors and health applications for measuring the user's heart rate, stress index, oxygen saturation, sleeping hours, step count, and so on. Wearable devices and several applications check these and additional user health conditions. The collected individual health records are sent to some form of personal healthcare manager through health applications via the smart devices. The user receives a feedback based on manager analysis. Therefore, the general public has a more accessibility to

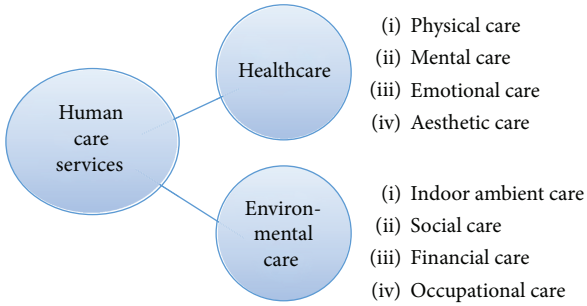


Figure 1: Types of human care services.

experience the e-healthcare services with the development of the related technology.

However, healthcare services are not limited to medical treatment. They can be extended to environmental management. Ulrich [4] proposed that psychological and social needs influence the medical outcomes of patients. He classified several environmental attributes including noise, light, windows, flooring materials, furniture arrangement, and air quality. These factors influence human sense, recognition, and emotion. Ulrich found that the environment can have a positive or a negative effect on patient recovery. This finding demonstrates the importance of ambient condition management. Since patients spend more time inside a building than outside, indoor environment could encourage patient recovery.

As the demand for quality of life increases, the wellness of individuals becomes more important than ever before. Thus, Figure 1 depicts how the scope and the range of human care services extend to incorporate the general public, not only for healthcare services, including physical, mental, emotional, and aesthetic cares, but also for environmental care, such as indoor ambient care, and social, financial, and occupational cares. Human care services can be applied to the ambient condition management of home, office, factory, and farm environments using a smartphone and other connected smart technology. For example, the ambient condition of a building, including lighting, temperature, and humidity, could be controlled by a smartphone that integrates IoT technologies. Recently, smart appliances have been linked to users and operated using signals sent from a smartphone or an unmanned repeater.

To improve ambient conditions for both patients and the general publics, multisensing and communication infrastructures are required. With the development of Internet-based technologies, it is easy to connect the related systems and individuals that are necessary to control the ambient building conditions. Indoor ambient care improves occupant comfort and convenience. Since the environmental conditions of a building are the product of interactions between the building's outdoor and indoor environments and subsystems, it is crucial to understand building physics in addition to occupant preferences. There have been many studies on using thermal analysis of buildings' inform design of heating, ventilation, and air-conditioning (HVAC) systems to predict energy consumption and to improve energy performance of the buildings [5–7]. However, there are few studies that concentrate on occupant activity as a meaningful influence on the thermal comfort of building occupants.

This study focuses on improving human care services provided by smart buildings as measured by increased occupant satisfaction with thermal conditions through the incorporation of considerations for occupant activity level in the analytical model. To accomplish this, Section 2 introduces the concept of the IoT platform of a smart building, which provides human care services including occupant comfort. Then, in Section 3, we suggest a thermal model of occupants and a control logic based on predicted mean vote (PMV). In Section 4, we integrate the proposed thermal comfort-based control logic into building models and obtain the indoor temperatures and the PMV values during different occupant activities. We compare the results of a PMV index-controlled heating system with those of a system thermally controlled by temperature. Finally, Section 5 relates our findings and conclusions.

2. IoT Platform of a Smart Building

A smart building is composed of automated building equipment and a communication infrastructure. The equipment includes HVAC systems, lighting systems, shading systems, window opening systems, elevators, air quality control systems, and other electrical devices and applications. Such dedicated equipment, categorized by functions, has been integrated with the smart building platform to facilitate real-time monitoring and controls using advanced technologies [8, 9]. However, these systems are not directly connected to each other due to the different communication protocols specified by their manufacturers. To solve this communication problem, there have been many attempts to standardize the protocols and to integrate them on the same platform.

A smart building is connected to users and rapidly replies to instantaneous demands of the users. Figure 2 shows the basic three-layer architecture of a smart building IoT platform including data sensing, data processing, and data reproduction. As depicted in Figure 2, the first layer covers the data collection of sensors. The data are composed of the individual conditions of users, operation states of appliances and equipment, and indoor ambient conditions such as temperature, illuminance, and humidity. These are transferred via an Internet-connected gateway and stored in a dedicated big data cloud. The next layer is the data processing. Data is classified and processed for the controlling actuators of the building. Comfort-related actuators such as the HVAC system, lighting system, and blind system require data from individual users. In the reproducing layer, classified and processed data are reproduced as information pertaining to each interaction between occupant and appliance. The data is accumulated in time series. The accumulated data becomes knowledge of the users and informs system efficacy and efficiency enhancements, providing better services for the occupants.

As stated in the previous section, building conditions are the product of building interactions with their environments and subsystems. To control actuators with a focus on user comfort requires additional data processing that considers user activity and its influence on building conditions.

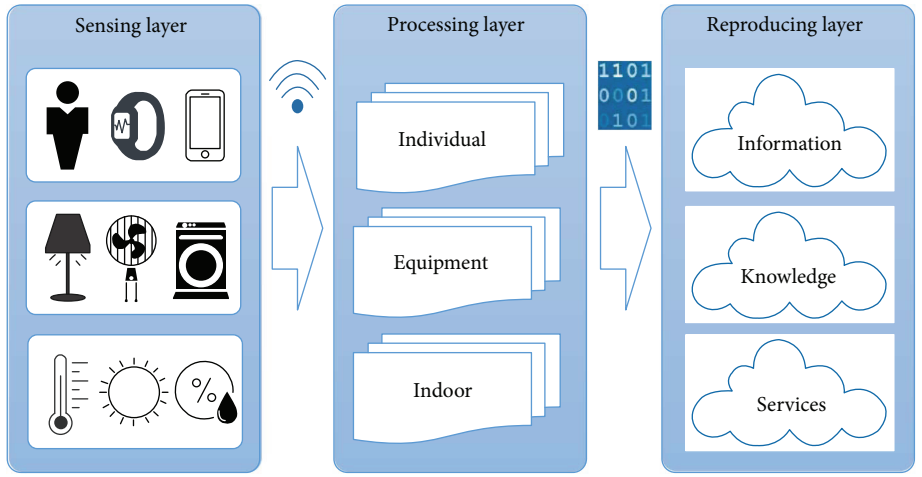


Figure 2: The basic architecture of an IoT platform in a smart building.

Table 1: Metabolic rates by activity.

Class	Mean of metabolic heat (W/m^2)	Metabolic heat (W)	Example
Resting	65	115	Resting
Low	100	180	Sitting at ease/standing
Moderate	165	295	Sustained hand/arm work
High	230	415	Intense work
Very high	290	520	Very intense to maximum activity

The following section proposes a thermal model of occupants that will improve human care services in IoT-based smart buildings.

3. Thermal Comfort of Occupants

3.1. Thermal Model of Occupants. The first step in conducting the thermal analysis of a building is to determine the internal heat gain of the building. Building internal heat gain is comprised of solar gain, metabolic heat gain, and heat gain produced by appliances. Metabolic heat gain, the heat gain profile of occupants, is a result of an occupant activity, such as resting, standing, and working. A heat gain model of occupants is deduced from the heat balance equation of the human body. The heat exchange of the human body in indoor conditions is expressed by the following equation [10, 11]:

$$M = S + \mathrm{RL} + \mathrm{CL} + \mathrm{EL}, \tag{1}$$

where M is the metabolic rate of generation of heat in the body (W/m^2); S is the storage or the rate of net loss of heat due to lowering of body temperature (W/m^2), counted negative when the body gains heat; RL is the rate of radiative loss of heat to the environment (W/m^2), negative when the walls or other radiative surfaces are warmer than the skin; CL is the rate of convective loss of heat to the environment (W/m^2), negative when the air is warmer than the skin; and EL is the rate of loss of heat by evaporation in the lungs and from the skin (W/m^2).

Metabolic heat production expresses the rate of production of energy with time. Hence, M represents the units of power (W). Since this term is related to surface area of the body, its unit is generally expressed by W/m^2. Moreover, the unit "MET" is sometimes used. "1 MET" is equivalent to $50\,kcal/m^2/h = 58.2\,W/m^2$ and is said to be the metabolic rate of a seated person at rest [12]. The ISO 8996 standard [13] gives data for estimating the metabolic heat production of a human body. Table 1 shows the classification of metabolic rates by activity. It provides the fundamental support to ISO thermal comfort and other standards.

A static model is used as a conventional thermal model of occupants. In this model, the most important parameters are the number of occupants and their heat gains expressed as time series. However, this model does not reflect the fact that the metabolic heat flux varies over time, depending on both the occupants and their environment. Moreover, heat is transferred by multiple mechanisms, including conduction, convection, and radiation from body to environment. Therefore, we propose a dynamic thermal model of occupants to provide more accurate analysis and improve services. The heat transfer through body-clothing-environment can be expressed as follows:

$$\phi_{\mathrm{core}} = \phi_{\mathrm{storage}} + \phi_{\mathrm{evaporation}} + \phi_{\mathrm{dissipation}}, \tag{2}$$

where ϕ_{core} is the metabolic heat gain of human body (W), ϕ_{storage} is the heat flux (W) for rising body temperature or stored heat flux within body, $\phi_{\mathrm{evaporation}}$ is the heat flux (W)

evaporated in the lungs and from the skin, and $\phi_{\text{dissipation}}$ is the dissipated heat flux (W) from body to environment. The body-clothing-environment heat transfer directly influences the thermal condition of environment, and this process is expressed as follows:

$$\phi_{\text{dissipation}} = C_{\text{body}} \frac{dT_{\text{body}}}{dt} + \frac{1}{R_{\text{body}}} \left(T_{\text{body}} - T_{\text{indoor}}\right), \quad (3)$$

where C_{body} is the thermal capacitance (J/K) of the human body. R_{body} is the thermal resistance (°C/W) of the body. T_{body} is the temperature of the body. T_{indoor} is the temperature (°C) of indoor environment. This is a first-order *RC* thermal network model. For more detail, this model can be developed into a second-order model, as shown below:

$$\begin{aligned}
\phi_{\text{dissipation}} &= \phi_{\text{body-clothing}} + \phi_{\text{clothing-indoor}}, \\
\phi_{\text{body-clothing}} &= C_{\text{body}} \frac{dT_{\text{body}}}{dt} + \frac{1}{R_{\text{body}}} \left(T_{\text{body}} - T_{\text{clothing}}\right), \\
\phi_{\text{clothging-indoor}} &= C_{\text{clothing}} \frac{dT_{\text{clothing}}}{dt} \\
&\quad + \frac{1}{R_{\text{clothing}}} \left(T_{\text{clothing}} - T_{\text{indoor}}\right),
\end{aligned} \quad (4)$$

where $\phi_{\text{body-clothing}}$ is the heat flux from body to clothing, and $\phi_{\text{clothing-indoor}}$ is the heat flux from clothing to indoor environment. C_{clothing} and R_{clothing} are the thermal capacitance (J/K) and the thermal resistance (°C/W) of the clothing of occupant, respectively. T_{clothing} is the temperature (°C) of clothing of the body.

The proposed model implies thermal characteristics of building occupants, which informs an understanding of the impact of individual thermal dynamics of building occupants. Factoring in the thermal dynamics of occupants enables a more accurate analysis of the thermal performance of a building. If this analysis is applied to the thermal control of a building, it can increase occupant comfort in terms of temperature conditions.

3.2. Thermal Comfort. Thermal comfort is an important indicator of overall building performance. It is defined as "that expression of mind which expresses satisfaction with the thermal environment" by the American Society of Heating, Refrigerating, and Air-Conditioning (ASHRAE) [14]. Since thermal comfort is personally determined and differs substantially between persons, it is not easy to quantify and analyze the value. Many researchers have investigated the parameters influencing thermal comfort in attempts to identify thermal comfort zones acceptable to the greatest number of people [15, 16].

The PMV model developed by Fanger in 1970s is the most well-known thermal comfort model. The PMV model is still applied to HVAC designs and referenced in recent studies [17–19]. To determine thermal comfort, this model evaluates six parameters: indoor air temperature, mean radiant temperature, relative humidity, air velocity, clothing, and metabolic rate of the occupant. The thermal comfort index is obtained as follows [20]:

$$\text{PMV} = \left[0.303 \cdot \exp(-0.036 \cdot M) + 0.028\right] \cdot L, \quad (5)$$

where

$$\begin{aligned}
L &= (M - W) - 3.05 \cdot 10^{-3} \cdot [5733 - 6.99 \cdot (M - W) - P_{\text{a}}] - 0.42 \cdot [(M - W) - 58.15] - 1.7 \cdot 10^{-5} \cdot M \cdot (5867 - P_{\text{a}}) \\
&\quad - 0.0014 \cdot M \cdot (34 - t_{\text{a}}) - 3.96 \cdot 10^{-8} \cdot f_{\text{cl}} \cdot \left[(t_{\text{cl}} + 273)^4 - (\bar{t}_{\text{r}} + 273)^4\right] - f_{\text{cl}} \cdot h_{\text{c}} \cdot (t_{\text{cl}} - t_{\text{a}}), \\
t_{\text{cl}} &= 35.7 - 0.028 \cdot (M - W) - I_{\text{cl}} \cdot 3.96 \cdot 10^{-8} \cdot f_{\text{cl}} \cdot f_{\text{cl}} \cdot \left[(t_{\text{cl}} + 273)^4 - (\bar{t}_{\text{r}} + 273)^4\right] - I_{\text{cl}} \cdot f_{\text{cl}} \cdot h_{\text{c}} \cdot (t_{\text{cl}} - t_{\text{a}}), \\
h_{\text{c}} &= \begin{cases} 2.38 \cdot |t_{\text{cl}} - t_{\text{a}}|^{0.25}, & \text{if } 2.38 \cdot |t_{\text{cl}} - t_{\text{a}}|^{0.25} > 12.1\sqrt{v_{\text{ar}}}, \\ 12.1\sqrt{v_{\text{ar}}}, & \text{if } 2.38 \cdot |t_{\text{cl}} - t_{\text{a}}|^{0.25} \le 12.1\sqrt{v_{\text{ar}}}, \end{cases} \\
f_{\text{cl}} &= \begin{cases} 1 + 1.29 \cdot I_{\text{cl}}, & \text{if } I_{\text{cl}} \le 0.078, \\ 1.05 + 0.645 \cdot I_{\text{cl}}, & \text{if } I_{\text{cl}} > 0.078, \end{cases} \\
P_{\text{a}} &= h_{\text{r}} \cdot 6.1094 \cdot \exp\left[\frac{(17.625 \cdot t_{\text{a}})}{(t_{\text{a}} + 243.04)}\right],
\end{aligned} \quad (6)$$

where PMV is the value of the predicted mean vote index, M is the metabolic rate (W/m^2), L is the thermal load of human body (W/m^2), W is the rate of mechanical work (W/m^2) which is 0 in most activities, P_{a} is the partial water vapor pressure (Pa), t_{a} is the indoor air temperature (°C), t_{cl} is the surface temperature of clothing (°C), $\bar{t}_{\text{r}}$ is the mean radiant

Table 2: PMV index.

Index	Explanation
−3	Cold
−2	Cool
−1	Slightly cool
0	Neutral
1	Slightly warm
2	Warm
3	Hot

temperature (°C), f_{cl} is the clothing surface area factor, I_{cl} is the thermal resistance of clothing ($m^{2\circ}C/W$), h_c is the convective heat transfer coefficient ($W/m^{2\circ}C$), v_{ar} is the air velocity (m/s), and h_r is the relative humidity (%). The PMV index is represented by 7 points from −3 to 3 as summarized in Table 2. The optimal temperature is achieved when PMV is zero, indicating thermally neutral sensation, during different human activity level [21].

3.3. PMV-Based Control Algorithm. Thermal condition of the indoor environment is determined by maintaining a PMV in the range of −0.2 to 0.2. This range is one of the recommended thermal environments given by ISO EN 7730 and CEN standard EN 15251 [22–24]. These standards suggested the PMV value between −0.2 and 0.2 for a high level of expectation and for spaces occupied by very sensitive and fragile persons. Since our study is focused on human care services including a healthcare service, we should consider the highest limitations on thermal comfort. It is why we limited the range of PMV from −0.2 to 0.2. The six parameters of PMV are measured by sensors installed in a target space and by personal wearable devices.

The proposed PMV-based control algorithm consists of following sequence of events. First, the presence of an occupant is detected to judge the necessity of operation of heating/cooling systems. If an occupant is detected inside a target space, the PMV index is calculated by using (5). In this step, we use metabolic heat gain of a human body obtained by the proposed thermal model of occupants. Since the thermal conditions of the space are controlled by the PMV index, by extension, the control command signals used to operate thermal systems such as heating and air-conditioning are also determined by the PMV value. For example, during winter season, a heating system operates. If the PMV value is less than −0.2, a command signal for a heating system is given as "1." Else, the command signal is given as "0." During summer season, a cooling system is used. If the PMV value is greater than 0.2, a command signal to operate the cooling system is given as "1." Otherwise, the command is "0." A PMV-based control is expected to make occupants more comfortable and satisfied than a temperature-based control.

4. Case Study

To investigate the feasibility of thermal modeling of occupants and the influence of metabolic heat gain inside a building, we integrated the proposed thermal model of occupants with a simple *RC*-lumped building model using MATLAB/Simulink. Then, the metabolic heat gain obtained by the proposed thermal model, which considered different occupant activities, was applied to the building model as a heat source. The thermal conditions of the building, as controlled by temperature and by the PMV index, were compared to determine which control algorithm is more useful for providing human care services in smart buildings.

Table 3 describes the parameters of a building model. These parameters were used in the modelling process and the simulation of our case study. As the weather condition, we selected a cold weather. The outdoor temperature is varied against time and is given as a sinusoidal function of which amplitude is 3. The temperature varies from −3 to 3. In addition, we considered the case where the heating system is required to achieve a certain range of indoor temperature or PMV values.

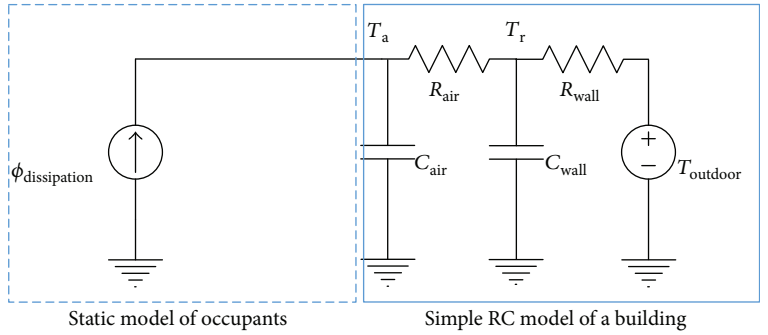
4.1. Metabolic Heat Gain of Occupants. To demonstrate how metabolic heat affects the indoor temperature of a building, we created a simple scenario of activities for a building occupant. The occupant stays in the building in a resting mode for most of the daily 24-hour period, dissipating metabolic heat of 115 W. For two hours and forty minutes, from 16:00 to 18:40, the occupant works with their hands and arms. This activity level increases dissipated metabolic heat gain to 295 W. We applied this scenario to the proposed thermal model of occupants. We developed two models in MATLAB/Simulink: (1) a static model and (2) a dynamic model. Figure 3 shows the developed thermal models of occupants integrated with a building model. The first model is a conventional model used for thermal analysis of buildings. It does not account for the thermal dynamics of users. The second model is the dynamic model proposed in Section 3. To describe the thermal characteristics of the human body, we used global thermal resistance and global thermal capacitance values of a human body as 30 W/m·K and 3770 J/kg/K, respectively.

After developing these models, we compared the influence of the heat gain generated by a static model and a dynamic model following the given occupant activity scenario inside the building. Figure 4 shows the simulated static and dynamic heat flux ($\phi_{dissipation}$) dissipated by the occupant due to their specific activity. The heat gain simulated with a static model follows a constant value for each activity. In a resting mode, the metabolic heat gain of 115 W is dissipated by the occupant. While the occupant sustainably works with their arms and hands, the heat of 295 W is dissipated. The static model only shows a heat gain profile of the occupant in a steady state. However, the dynamic model describes a heat gain profile in a steady state and a transient state since this model considers thermal resistance and capacitance values of the occupant. In this model, heat is charged and discharged according to thermal characteristics of the occupant.

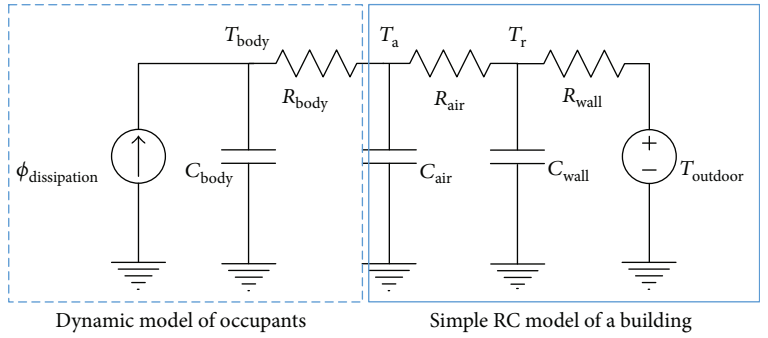
Since the occupant is thermally linked to the building, the activity of the occupant should be considered in the management of the thermal comfort and energy use of the building. Given the generated metabolic heat flux, the temperature

Table 3: A brief description of the used building.

Parameters	Description
Building type	Residential building
Floor area	Modular space area = 52.8 m^2
Dimension and heights	16 m × 3.3 m; floor-to-ceiling = 4 m
Wall	Thickness = 0.2 m; density = 1920 kg/m^3; specific heat = 835 J/kg/K; thermal conductivity = 0.038 W/m/K
Window	Thickness = 0.01 m; density = 2700 kg/m^3; specific heat = 840 J/kg/K; thermal conductivity = 0.78 W/m/K
Operating hours	24 hours
Metabolic heat gain	Resting mode: 115 W
Outdoor temperature	−3°C~3°C



(a)



(b)

Figure 3: Thermal models of occupants integrated in a building model: (a) static model and (b) dynamic model.

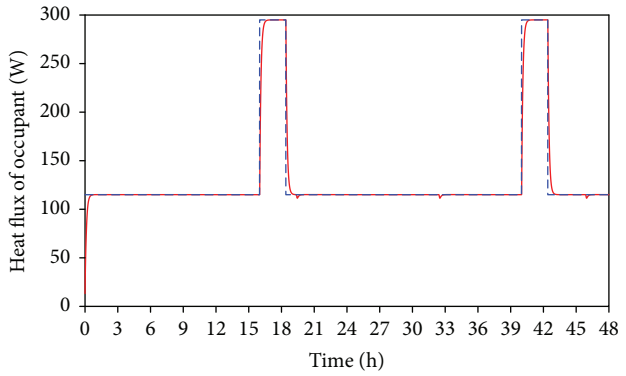


Figure 4: Heat flux of the occupant (blue broken line: static model; red line: dynamic model).

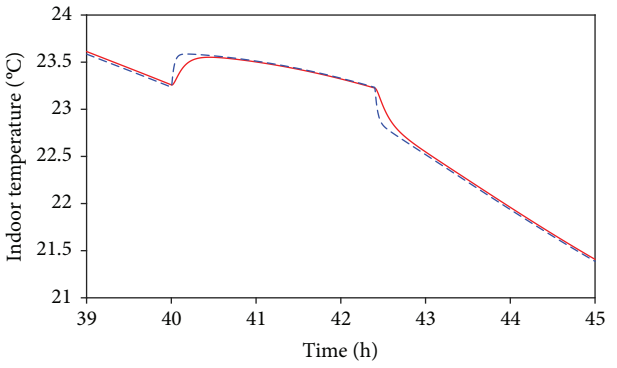


Figure 5: Indoor temperature of the building (blue broken line: static model; red line: dynamic model).

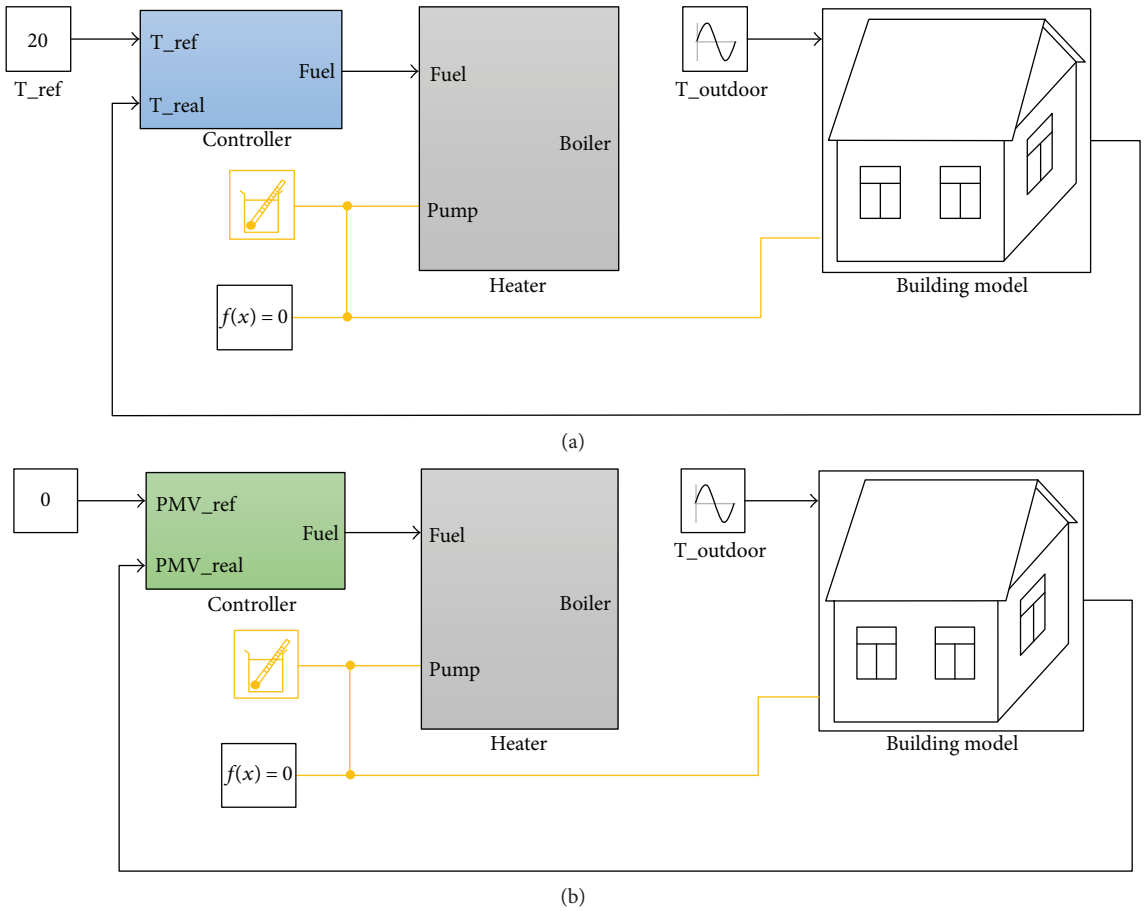


Figure 6: Important features of the building and its subsystems: (a) temperature-based controller and (b) PMV-based controller.

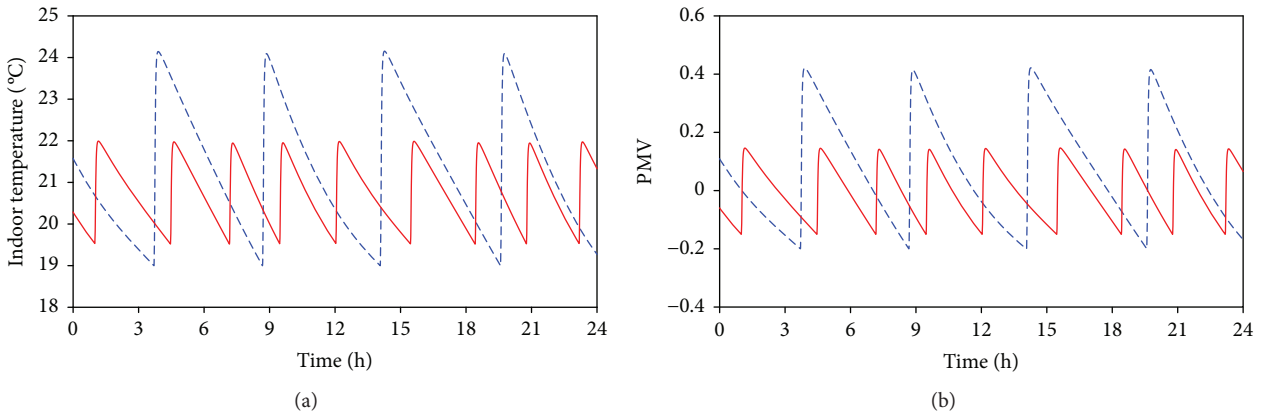
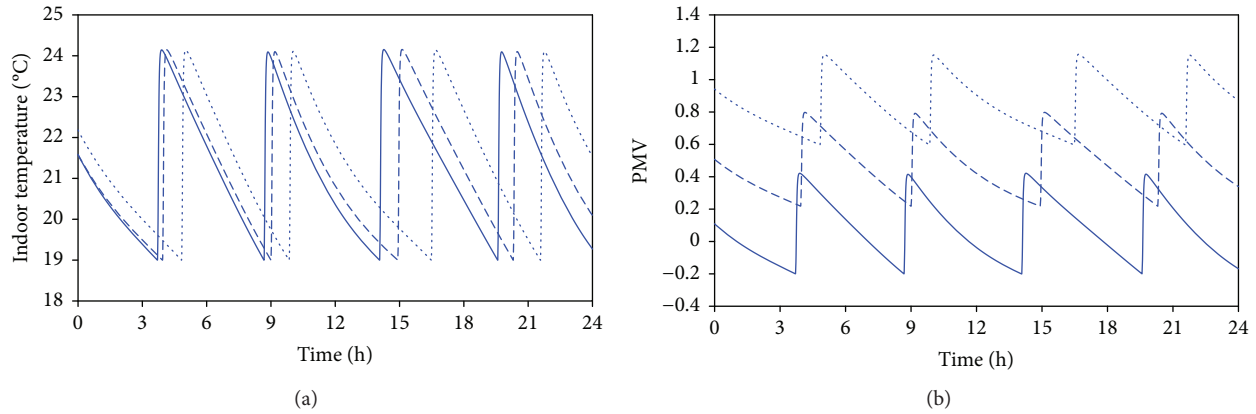


Figure 7: Temperature and PMV: (a) T_a and (b) PMV (blue broken line: model 1; red line: model 2).

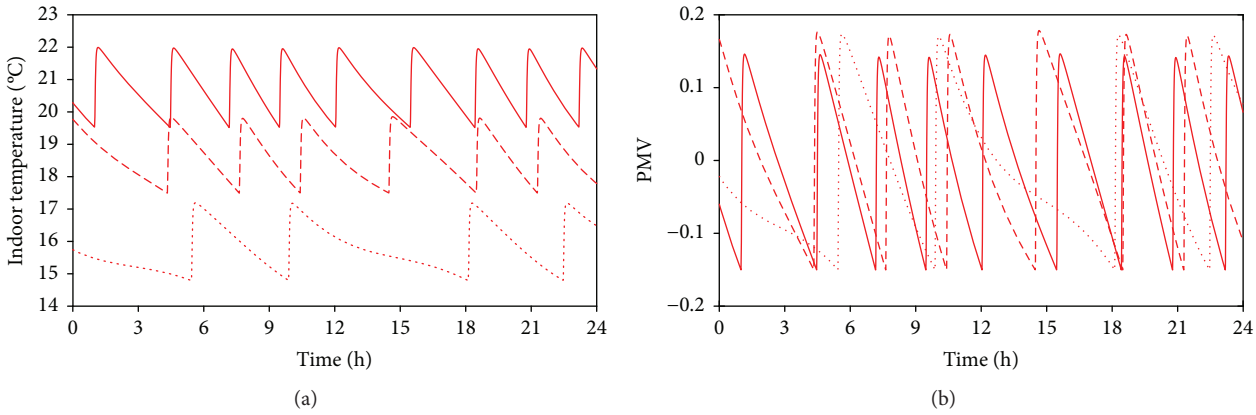
evolution of the building is obtained as shown in Figure 5. The indoor temperature of the building (T_a) is influenced by the occupant, demonstrating that the occupant can be considered a heat source of the building. The indoor temperature globally decreases during in the period from hour 39 to hour 45. This period is the equivalent of hour 15 to hour 21 of the second 24-hour cycle.

Despite the global decrease, note the significant temperature variation during the period from 40:00 to 42:40 or 16:00 to 18:40 of the second daily period, of the simulation. In the scenario, this is the period when the occupant continuously works with their arms and hands. While T_a obtained from the static model shows the thermal dynamics of the building, T_a given by the dynamic model depicts the dynamics of the occupant as well as that of the building. This demonstrates that the dynamic model describes a more accurate temperature evolution of the building. Consequently, the proposed model performs better, providing a more detailed thermal analysis, enabling an improved control strategy for thermal systems and increasing the thermal comfort of occupants.



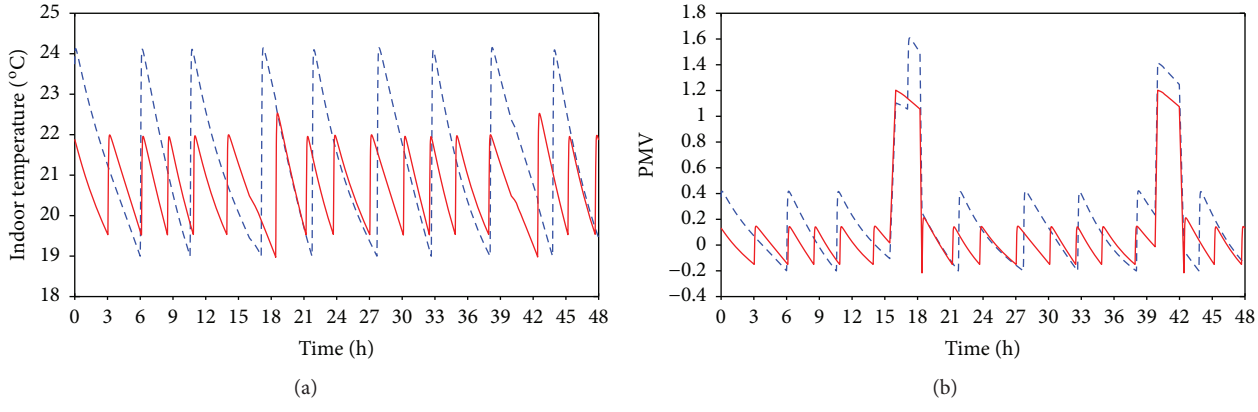
(a) (b)

Figure 8: Temperature and PMV of model 1 with different $\phi_{\text{dissipation}}$: (a) T_a and (b) PMV (blue line: 115 W; blue broken line: 135 W; blue dotted line: 155 W).



(a) (b)

Figure 9: Temperature and PMV of model 2 with different $\phi_{\text{dissipation}}$: (a) T_a and (b) PMV (red line: 115 W; red broken line: 135 W; red dotted line: 155 W).



(a) (b)

Figure 10: Temperature and PMV: (a) T_a and (b) PMV (blue broken line: model 1; red line: model 2).

Therefore, we suggest integrating this model with IoT-based smart building environment services for occupants' comfort.

4.2. PMV-Based Thermal Control. The general features of the smart building, heating system, and controller for each of the two models are illustrated in Figure 6. The first model has a heater with a temperature-based controller. Model 1 controls the heating system based on a reference value for indoor temperature. Model 2 achieves a PMV-led control for thermal comfort of building occupants. Model 2 maintains a PMV value in the range of −0.2 to 0.2 as explained in Section 3.3.

Thermal controls based on temperature and PMV are both achieved based on the scenario of an occupant inside the building. Figure 7 shows the results of T_a of the building and the PMV of an occupant in a resting mode. T_a of model 1 is controlled between 19°C and 24°C, and the value of PMV of model 1 follows the variation of T_a, ranging between −0.2 and 0.4. Model 2 bases control of the heater on the PMV index. If the PMV is smaller than its reference value, the heater is activated. Then, T_a increases. If the PMV value is greater than its reference value, the heater ceases operations. Consequently, T_a decreases. In both models, the indoor temperature and the PMV interact to each other because these are directly related to the operation of the heating system. Therefore, the trend of indoor temperature evolution is similar to that of PMV.

However, the heat dissipation of the occupant does not directly affect T_a in model 1 because the heater of model 1 operates in accordance with T_a as shown in Figure 8. Although the heat dissipation levels of the occupant become different as 115 W, 135 W, and 155 W, the range of T_a does not change. Since thermal condition of the building is controlled by temperature, it is reasonable to obtain a fixed range of T_a irrelevant to any activities of the occupant. Moreover, the considered metabolic heat gains of 115 W, 135 W, and 155 W are small enough to keep the temperature within the reference between 19°C and 24°C. It is why the temperature does not surpass 24°C and that there was no requirement for a cooling system. However, it would be required to integrate the cooling system for the case where heat dissipation of occupants is high enough.

In the context of the PMV index, stronger activities cause higher metabolic heat dissipation and leads to different levels of occupant thermal comfort. Therefore, with higher metabolic heat, the value of PMV becomes higher in the same temperature condition of the building. Consequently, thermal comfort of the occupant is not assured when applying model 1 with a temperature-based controller.

Contrary to model 1, occupant thermal comfort is preferentially assured in model 2 which implements PMV-based control logic. Figure 9 depicts the T_a and PMV obtained in model 2. While the range of T_a is fixed from 19°C to 24°C for all three activities in model 1, the range of T_a in model 2 shifts to account for the different heat dissipation levels of each activity. In model 2, the PMV range is fixed. To achieve PMV values within the optimum range of −0.2 and 0.2, the operation of the heater is controlled as specified in Section 3.3. Since greater heat dissipation induces the occupant to feel hotter, the range of T_a is lower than when less metabolic heat dissipation occurs. This is demonstrated by the observation that the higher the metabolic heat, the lower the T_a, while lower metabolic heat values result in higher T_a. Since the heater operates in accordance with the PMV of the occupant, the ranges of the obtained indoor temperature under different quantity of metabolic heat gain of the occupant are differently determined while the ranges of the PMV are similar to each other. Moreover, the indoor temperature decreases when the heater is turned off because the amplitude of the outdoor temperature is between −3 and 3°C, less than the indoor temperature.

We also implemented a scenario with different occupant activities in 24-hour cycles. The occupant rests during most of the daily cycle and dissipates metabolic heat of 115 W. However, from 16:00 to 18:40 of each 24 hour period, the occupant performs a low energy activity and dissipates metabolic heat of 180 W. Figure 10 presents the T_a and PMV values for the two models over the course of two consecutive 24-hour cycles of the previously described scenario. The impact of increased occupant heat dissipation is observed in the slight increase in T_a during occupant activity, despite a globally decreasing trend of T_a in both models. Analogously, the increased heat gain also raises T_a for both models under a globally increasing trend of T_a. Note that the PMV of model 1 peaks farther from the ideal, neutral zero point than does model 2. As discussed before, these results show that increased satisfaction with human care services is achieved when comfort-based actuators are controlled on the basis of comfort-based indices.

5. Conclusions

To increase comfort and convenience, the scope of human care services has expanded to include ambient condition management. With the help of the sensing and communication technologies, individual occupant health profiles could be acquired and accumulated in an IoT-based smart building. The profiles would be analysed for occupant information, and the applied knowledge would enable improved services to be implemented. To achieve thermal comfort for building occupants, we proposed a dynamic thermal model of occupants based on the heat balance equation of human body and thermal characteristics of the occupants. We implemented this model in two smart building models with heaters controlled by temperature and by the PMV value, the most widely used thermal comfort index. The simulation results showed that PMV-based thermal control improves occupant thermal satisfaction when compared to temperature-based thermal control. Therefore, we suggest that PMV-based thermal control be integrated into the IoT-based smart building platform to enable improved human care services.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2017R1D1A3B03035693). In addition, this work was supported by the Brain Korea 21 Plus Program (22A2015 2113298) funded by the National Research Foundation of Korea (NRF).

References

[1] S. Tang, V. Kalavally, K. Y. Ng, and J. Parkkinen, "Development of a prototype smart home intelligent lighting control architecture using sensors onboard a mobile computing system," *Energy and Buildings*, vol. 138, pp. 368–376, 2017.

[2] Y. Liu and X. Xu, "Industry 4.0 and cloud manufacturing: a comparative analysis," *Journal of Manufacturing Science and Engineering*, vol. 139, no. 3, article 034701, p. 8, 2017.

[3] K. Saleem, A. Derhab, J. Al-Muhtadi, and B. Shahzad, "Human-oriented design of secure machine-to-machine communication system for e-healthcare society," *Computers in Human Behavior*, vol. 51, pp. 977–985, 2015.

[4] R. S. Ulrich, "Effects of healthcare environmental design on medical outcomes," in *International Academy for Design and Health: Proceedings of the Second International Conference on Health and Design*, pp. 49–59, Sweden: Svensk Byggtjanst, 2001.

[5] A. T. Nguyen, S. Reiter, and P. Rigo, "A review on simulation-based optimization methods applied to building performance analysis," *Applied Energy*, vol. 113, pp. 1043–1058, 2014.

[6] W. I. W. M. Nazi, M. Royapoor, Y. Wang, and A. P. Roskilly, "Office building cooling load reduction using thermal analysis method – a case study," *Applied Energy*, vol. 185, pp. 1574–1584, 2017.

[7] S. Heinen, W. Turner, L. Cradden, F. McDermott, and M. O'Malley, "Electrification of residential space heating considering coincidental weather events and building thermal inertia: a system-wide planning analysis," *Energy*, vol. 127, pp. 136–154, 2017.

[8] C. Fan and F. Xiao, "Assessment of building operational performance using data mining techniques: a case study," *Energy Procedia*, vol. 111, pp. 1070–1078, 2017.

[9] N. Aste, M. Manfren, and G. Marenzi, "Building automation and control systems and performance optimization: a framework for analysis," *Renewable and Sustainable Energy Reviews*, vol. 75, pp. 313–330, 2017.

[10] D. Brunt, "Some physical aspects of the heat balance of the human body," *Proceedings of the Physical Society*, vol. 59, no. 5, pp. 713–726, 1947.

[11] H. Park, "Dynamic thermal modeling of electrical appliances for energy management of low energy buildings," in *Electric Power*, Université de Cergy-Pontoise, Cergy Pontoise, France, 2013.

[12] G. Havenith, I. Holmér, and K. Parsons, "Personal factors in thermal comfort assessment: clothing properties and metabolic heat production," *Energy and Buildings*, vol. 34, no. 6, pp. 581–591, 2002.

[13] ISO 8996, *Ergonomics of the Thermal Environments-Determination of Metabolic Heat Production*, ISO, Geneva, 1989.

[14] ASHRAE 55, *Thermal Environmental Conditions for Human Occupancy*, American Society of Heating Refrigerating Air-Conditioning Engineers, Atlanta, GA, USA, 2004.

[15] J. Hensen, "On the thermal interaction of building structure and heating and ventilation system, [Ph.D. Thesis]," Technische Universiteit Eindhoven, Eindhoven, Netherlands, 1991.

[16] H. Zhang, "Human thermal sensation and comfort in transient and non-uniform thermal environments, [Ph.D. Thesis]," University of California, Berkeley, CA, USA, 2003.

[17] L. Zampetti, M. Arnesano, and G. M. Revel, "Experimental testing of a system for the energy-efficient sub-zonal heating management in indoor environments based on PMV," *Energy and Buildings*, vol. 166, pp. 229–238, 2018.

[18] Z. Xu, G. Hu, C. J. Spanos, and S. Schiavon, "PMV-based event-triggered mechanism for building energy management under uncertainties," *Energy and Buildings*, vol. 152, pp. 73–85, 2017.

[19] R. Holopainen, P. Tuomaala, P. Hernandez, T. Hakkinen, K. Piira, and J. Piippo, "Comfort assessment in the context of sustainable buildings: comparison of simplified and detailed human thermal sensation methods," *Building and Environment*, vol. 71, pp. 60–70, 2014.

[20] P. Fanger, *Thermal Comfort: Analysis and Applications in Environmental Engineering*, Danish Technical Press, Copenhagen, 1970.

[21] ISO7730, *Moderate Thermal Environments – Determination of the PMV and PPD Indices and Specification of the Conditions for Thermal Comfort*, ISO, Geneva, 1984.

[22] ISO7730, *Ergonomics of the Thermal Environment Analytical Determination and Interpretation of Thermal Comfort Using Calculation of the PMV and PPD Indices and Local Thermal Comfort Criteria*, ISO, Geneva, 2005.

[23] F. Nicol and M. Wilson, "An overview of the European standard EN 15251," in *Proceedings on Adapting to Change: New Thinking Comfort*, pp. 1–13, Cumberland Lodge, Windsor, UK, April 2010.

[24] EN15251, *Indoor Environmental Input Parameters for Design and Assessment of Energy Performance of Buildings – Addressing Indoor Air Quality, Thermal Environment, Lighting and Acoustics*, European Committee for Standardization, Brussels, 2007.

Hindawi
Journal of Sensors
Volume 2018, Article ID 3176967, 9 pages
https://doi.org/10.1155/2018/3176967

Research Article

Efficient Heterogeneous Network-Routing Method Based on Dynamic Control Middleware for Cyber-Physical System

Hyungsoo Lee, Jaehwan Lee, Sanghyuck Nam, and Sangoh Park

School of Computer Science and Engineering, Chung-Ang University, Seoul 06974, Republic of Korea

Correspondence should be addressed to Sangoh Park; sopark@cau.ac.kr

Received 22 February 2018; Accepted 23 April 2018; Published 20 May 2018

Academic Editor: Ka L. Man

A cyber-physical system depends on stable control and interaction between the many systems and devices connected to the network. Dynamic control middleware, which considers the characteristics of a cyber-physical system, supports the dynamic search and control of devices existing on the global network using Internet protocol version 6 (IPv6). However, systems and devices may connect to a network using a variety of heterogeneous protocols, not just IPv6. To solve the problem of heterogeneous protocols, this paper proposes a routing technique which enables network devices to communicate using different protocols. The proposed network-routing module can register devices with various protocols and improve the stability of the efficient heterogeneous network.

1. Introduction

Cyber-physical systems (CPS) [1–7] consider the physical characteristics of the embedded system. A CPS depends on stable control and interaction between the many systems and devices connected to the network. Thus, CPS enables the convergence of the varied components of information technology (IT) that are crucial to every industry, including energy, transportation, health care, and defense. Therefore, a CPS can facilitate the construction of a network that can contain a large area. Unlike existing middleware [8–10], dynamic control middleware (DCM) [4] efficiently constructs a large-scale network using an Internet protocol version 6- (IPv6-) based two-layer architecture that supports dynamic search and control of networked devices. However, real-world systems and devices connect to the network using numerous heterogeneous protocols rather than a single protocol. For example, the myriad of small devices employed in health care use a variety of protocols. However, existing DCM does not support these varied heterogeneous protocols. Therefore, to connect devices that use heterogeneous protocols, middleware must provide a routing method that enables various protocols to communicate with each other. Furthermore, health care devices are frequently connected and disconnected from the network owing to frequent movements. This paper suggests providing a dynamic reconfiguration method to stabilize the network by addressing the errors caused by frequent and unanticipated device connections and disconnections. In this paper, we propose extended middleware that can register heterogeneous devices as components of DCM using a variety of protocols such as IPv6, Bluetooth, and ZigBee. To improve the stability of the efficient heterogeneous network, this paper also suggests a mechanism to dynamically reconfigure the network even if the controller responsible for the intermediate node in the routing path is unexpectedly removed from the network due to unanticipated errors.

This paper is organized as follows. In the next section, we discuss related research. The Materials and Methods is comprised of three subsections: Subsection 3.1 introduces an extended DCM system with a novel network-routing module that supports various heterogeneous network protocols. Subsection 3.2 of Materials and Methods explains how to construct the routing table when inserting or deleting nodes. Subsection 3.3 introduces a testbed that creates a hybrid network environment which combines several heterogeneous

protocols. Here, we demonstrate that a message can be normally delivered. Section 4 presents conclusions and future research topics.

2. Related Work

Dynamic control middleware (DCM) efficiently controls connected devices throughout the IPv6-based global network, across all temporal and physical conditions. To search efficiently, DCM has a two-layer logical architecture consisting of a lower layer and an upper layer. The lower layer consists of control devices (CDs) connected to each local network, and the upper layer consists of control device managers (CDMs) that manage information from lower-layer devices. The upper layer can reduce the number of messages by only sending messages to devices belonging to the same IPv6 multicast group. In addition, because DCM has a two-layer architecture, and not a centralized server structure, devices can discover, search, and control other devices through multicasting, regardless of their global network location. However, DCM supports only IPv6 and in real-world applications, networked systems and devices often use different heterogeneous protocols.

Object-based middleware for smart home network (OSHNet) [11] supports interoperability regardless of device protocols using a device-routing table (DRT). However, although the DRT can deliver messages to other smart devices in the home, it cannot control external devices on the global network.

3. Materials and Methods

3.1. Extended DCM. Figure 1 diagrams the architecture of the proposed extended DCM (EDCM). The EDCM added a network-routing module so that CDs that do not support IPv6 can control each other with other devices. The network-routing module determines the path of the message and selects the appropriate network protocol for each transmission. The EDCM proposed in this paper can support various network protocols including IPv4, Bluetooth, ZigBee, and IPv6. In this subsection, we examine the four components that comprise the network-routing module.

3.1.1. Routing Mapper. The routing mapper identifies the destination of the message delivered from the upper layer of the EDCM. After referring to the routing table, the routing mapper maps the route to the message destination and adds it to the head of the message.

Table 1 shows an example of a routing table. The "Source" is the starting node. The "Destination" is the destination node. The "Interface" is the network protocol used to transmit messages between the two nodes. The "Source address" is the address of the starting node, and the "Destination address" is that of the destination node.

Box 1 shows an example of the content added by the routing mapper to the message head. Line 01 is the number of remaining routes to the destination node. Line 02 and Line 03 indicate the starting and destination nodes, while Line 04 through Line 24 store the mapping information in the routing table. Each time a message passes through an intermediate node, the number of routes in Line 01 is decremented by one, and one instance of routing table tag data is removed from the set represented between Line 04 and Line 24.

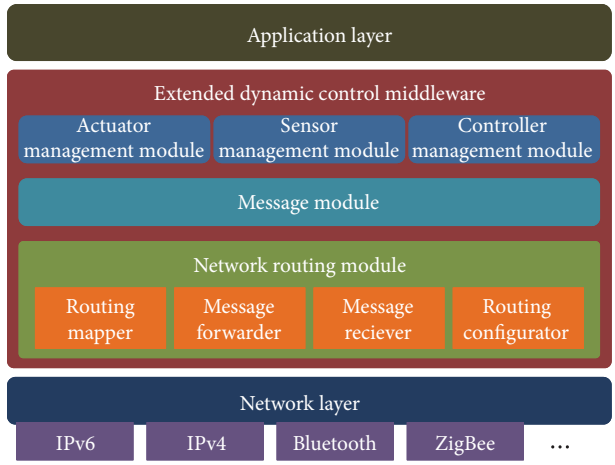


Figure 1: Extended DCM architecture.

Figure 2 shows a flowchart of the routing mapper. The routing mapper receives the message (MSG) from the upper layer. If the command of the MSG is "CDM Search Request," the routing mapper generates a "CDM Route Search Request" (CSRQ) message to obtain the route from the other CD to the CDM. Then, the message is forwarded to the message forwarder. At this time, the RouteCfg of the CSRQ message is changed to "true" to send the message to the routing configurator when the message arrives at another CD. The CD that sent the CSRQ message checks if there is routing information in the routing table to communicate with the CDM. If the required routing information is in the table, then the MSG is forwarded using the message forwarder.

3.1.2. Message Forwarder. Figure 3 shows a flowchart of the message forwarder. The message forwarder receives the MSG from the routing mapper or message receiver. The message forwarder extracts the NextRoute from the MSG and delivers the MSG based on the extracted information. At this time, the network protocol to be used is determined according to the information of the NextRoute.

If the message transmission fails, a new routing path is established using the CSRQ message to reconfigure the path currently connecting the CDMs. The message forwarder then retransmits the MSG using the modified routing information.

3.1.3. Message Receiver. Figure 4 shows a flowchart of the message receiver. The message receiver receives the MSG through various communication protocols such as IPv6, Bluetooth, or ZigBee. If MSG.Route is nonzero, the MSG is passed to the next path through the message forwarder. If MSG.Route is zero, the message receiver checks whether MSG.RouteCfg is "True" or "False." If "True," the MSG is a message related to the routing configuration and is forwarded to the routing configurator. If "false," the MSG is a generic message and is forwarded to the upper layer.

Table 1: Example of routing table.

Source	Destination	Interface	Source address	Destination address
CDM	CD-1	IPv6	2001:1F00:388::2	2001:1F00:388::3
CD-1	CD-2	Bluetooth	01:23:45:67:89:AB	01:23:45:67:89:AC
CD-2	CD-3	ZigBee	1001	1002

```
01: <Route>3</Route>
02: <Src>CDM</Src>
03: <Dest>CD-3</Dest>
04: <Routing_Table>
05:       <Src>CDM</Src>
06:       <Dest>CD-1</Dest>
07:       <Interface>IPv6<Interface>
08:             <Src_Address>2001:1F00:388::2</Src_Address>
09:             <Dest_Address>2001:1F00:388::3</Src_Address>
10: </Routing_Table>
11: <Routing_Table>
          <!– CD-1 to CD-2 –>
17: </Routing_Table>
18: <Routing_Table>
          <!– CD-2 to CD-3 –>
24: </Routing_Table>
          <!– End of Routing Table Part –>
```

Box 1: Routing table part in extended dynamic control middleware (ECDM).

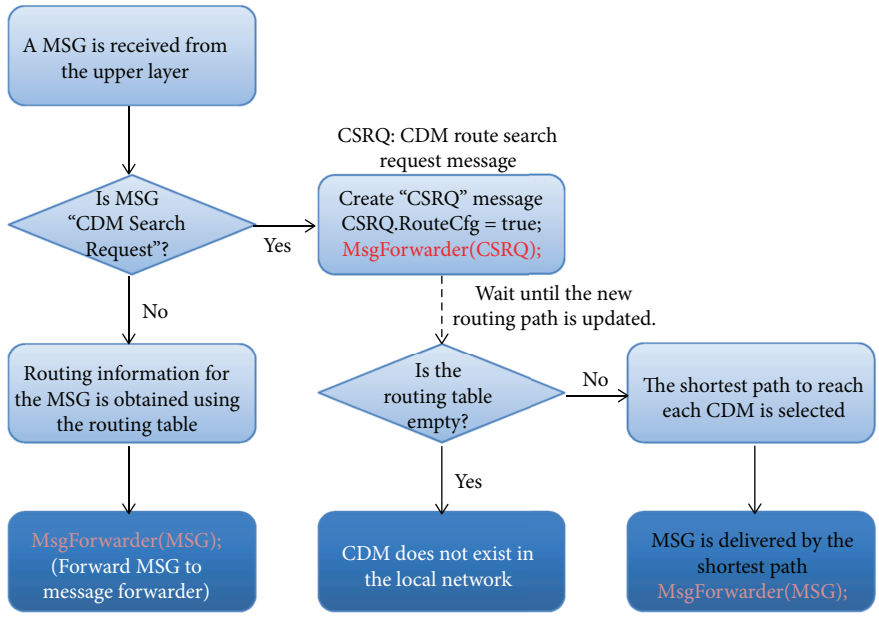


Figure 2: Flowchart of the routing mapper.

3.1.4. Routing Configurator. Figure 5 shows the flowchart of the routing configurator. The routing configurator that receives the MSG from the message receiver determines whether MSG is a CSRQ message. In the case of a CSRQ message, a "CDM Route Search Response" (CSRP) message is created. The CSRP message includes route information from the current controller to the CDM. In the case of a CSRP message, the routing configurator checks whether there is an existing CDM route. If no route exists, the route obtained from the CSRP message is stored in the routing table. If the routing table has an existing route, the length of the route obtained from the CSRP message is compared with that of the existing route. The shorter of the two compared routes is then stored or retained in the routing table.

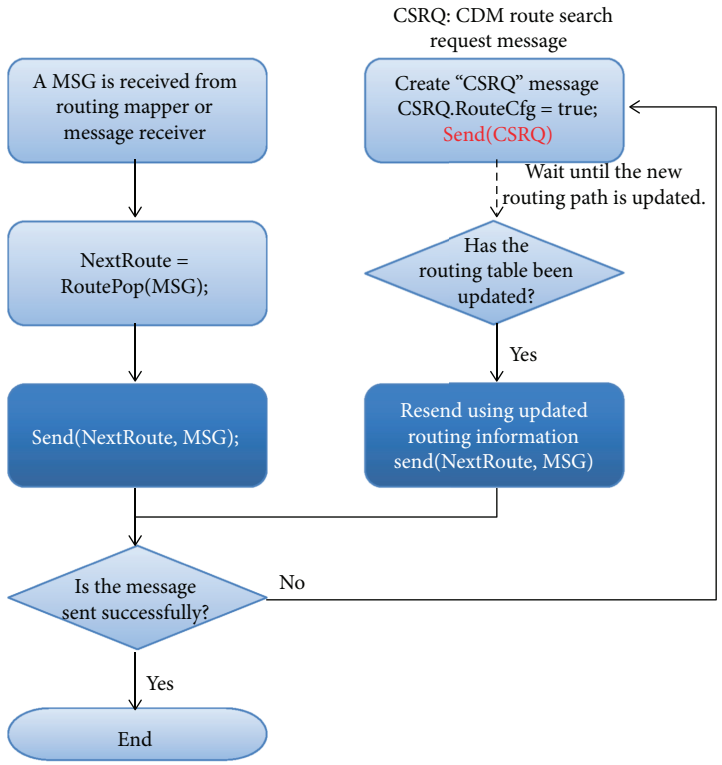


Figure 3: Flowchart of the message forwarder.

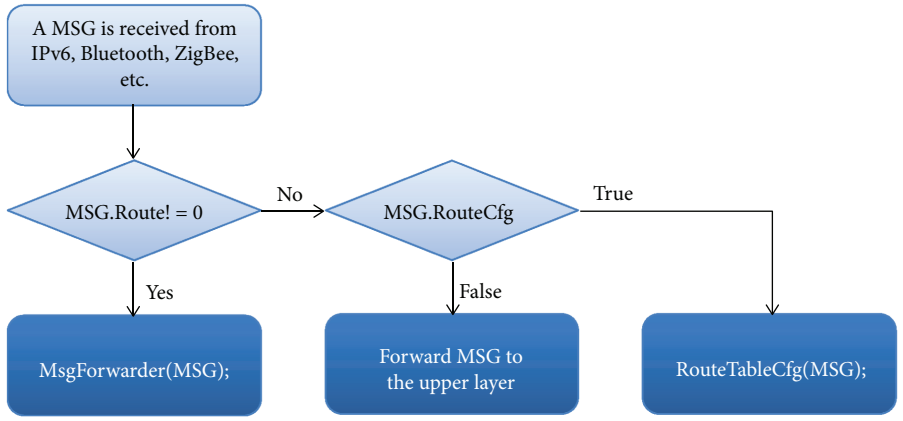


Figure 4: Flowchart of the message receiver.

3.2. Routing Table Configuration Policy. When a new CD is connected to a local network in the DCM, all CDs and CDMs use the same network protocol, IPv6, so the new CD sends a direct message to the CDM to register as a lower node of the CDM. However, although the DCM may support heterogeneous networks, the CDM may not support the specific network protocols used by the new CD.

Figure 6 shows an example of a CDM configuration in an EDCM regional network. Figure 6 assumes that CD-1 and CD-2, and CD-2 and CD-3, share physical proximity close enough to allow wireless transmission. The configuration also assumes that the CDM supports only IPv6, while CD-1 supports IPv6 and Bluetooth, CD-2 supports Bluetooth and ZigBee, and CD-3 only supports ZigBee. In this scenario, because CD-2 and CD-3 do not support IPv6, they cannot directly exchange messages with the CDM. Thus, CD-2 and CD-3 can only send and receive messages to and from the CDM by routing through one or more other CDs that directly or indirectly connect to the CDM. As shown Figure 6, CD-2 can indirectly connect to the CDM by using Bluetooth to connect to CD-1, which has a direct IPv6 connection to the CDM. Extending the concept of indirect connection, CD-3 can be indirectly connected to the CDM by connecting to CD-2 using ZigBee and then following the previously described connections from CD-2 to CD-1 to the CDM.

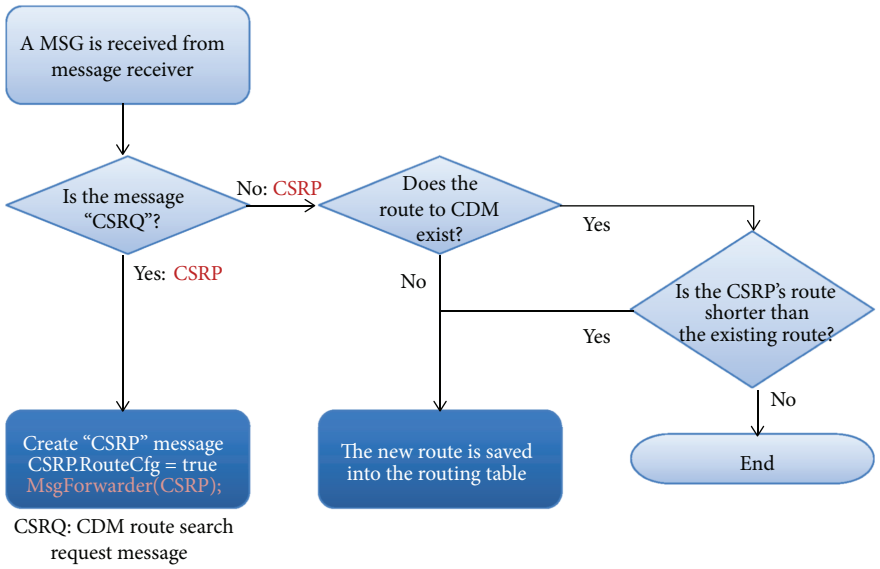


Figure 5: Flowchart of the routing configurator.

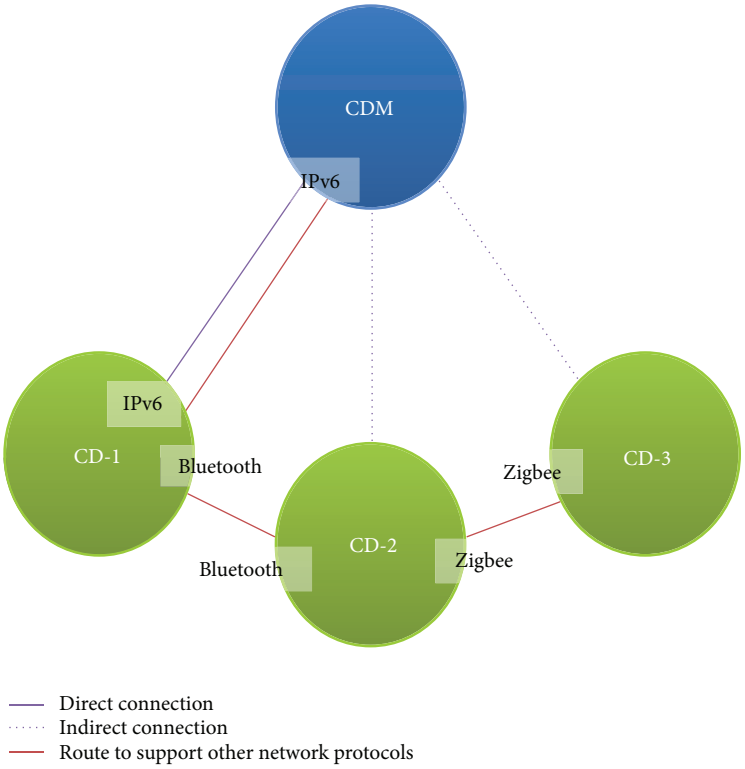


Figure 6: CDM configuration in ECDM local network.

Table 2: CD-1 routing table.

Source	Destination	Interface	Source address	Destination address
CDM	CD_1	IPv6	2001:1F00:388::2	2001:1F00:388::3

Tables 2–4 are the routing tables held by each CD. This routing table enables direct and indirect message transmission between CDM and CD.

The following subsections describe the network topology reconfiguration technique applied to solve the problems that occur when a new CD is connected and other connected CDs

Table 3: CD-2 routing table.

Source	Destination	Interface	Source address	Destination address
CDM	CD-1	IPv6	2001:1F00:388::2	2001:1F00:388::3
CD-1	CD-2	Bluetooth	01:23:45:67:89:AB	01:23:45:67:89:AC

Table 4: CD-3 routing table.

Source	Destination	Interface	Source address	Destination address
CDM	CD-1	IPv6	2001:1F00:388::2	2001:1F00:388::3
CD-1	CD-2	Bluetooth	01:23:45:67:89:AB	01:23:45:67:89:AC
CD-2	CD-3	ZigBee	1001	1002

Figure 7: New controller device connection in heterogeneous networks. (a) A new CD, CD-4, is connected to the EDCM. (b) CD-1 responds to CD-3. (c) A routing path is completed between CD-4 and the CDM.

Table 5: CD-4 routing table.

Source	Destination	Interface	Source address	Destination address
CDM	CD-1	IPv6	2001:1F00:388::2	2001:1F00:388::3
CD-1	CD-4	Bluetooth	01:23:45:67:89:AB	01:23:45:67:89:AD

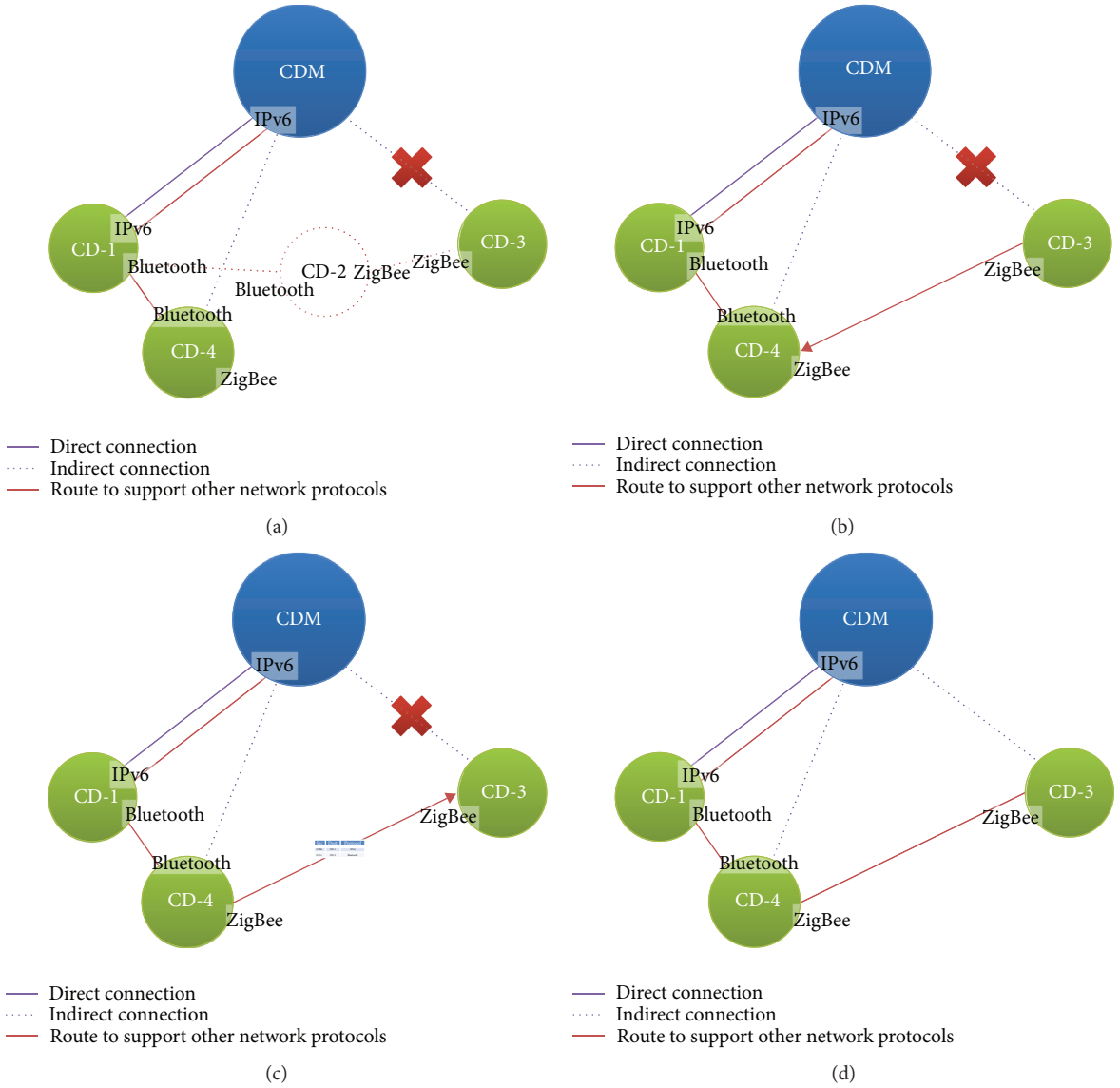


Figure 8: Dynamic reconfiguration in heterogeneous networks. (a) Intermediate node CD-2 is disconnected. (b) CD-3 searches for a new connection using broadcasting. (c) CD-4 transmits its routing table to CD-3. (d) CD-3 establishes a new routing path through CD-4.

are abnormally terminated. The proposed network topology reconfiguration technique is based on the CDM configuration shown in Figure 6.

3.2.1. New Controller Connection. Figure 7(a) depicts a new CD, CD-4, connecting to the EDCM local network. Since CD-4 does not support IPv6, CD-4 registers with the CDM by transmitting its information to the network indirectly through the other connected CDs using Bluetooth and ZigBee connections.

Figure 7(b) shows the responses of the other CDs to CD-4. After receiving the connection message of CD-4, CD-1, CD-2, and CD-3 all transmit their routing table information to CD-4. CD-4 receives the routing table information previously detailed in Tables 2–4 from all three CDs. If a message is sent over the routing path of CD-1, it reaches the CDM in two hops. However, if a message is transmitted through the routing paths through CD-2 or CD-3, two to four hops may be required. Therefore, CD-4 selects the shortest route, consisting of the fewest hops, and the routing table shown

Table 6: New routing table of CD-3.

Source	Destination	Interface	Source address	Destination address
CDM	CD-1	IPv6	2001:1F00:388::2	2001:1F00:388::3
CD-1	CD-4	Bluetooth	01:23:45:67:89:AB	01:23:45:67:89:AD
CD-4	CD-3	ZigBee	1003	1002

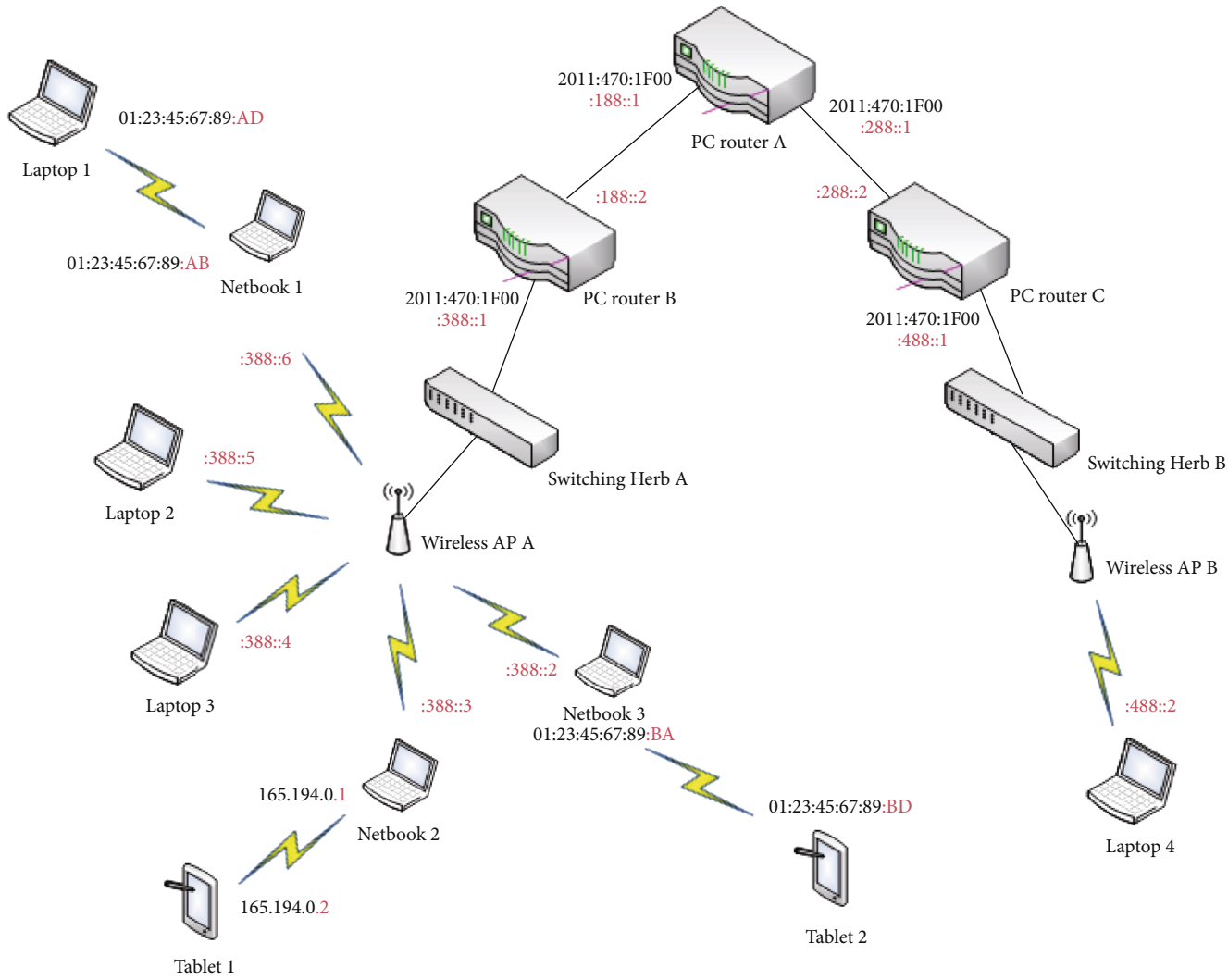


Figure 9: Testbed heterogeneous network-routing model.

in Table 5 is created for CD-4. Figure 7(c) shows that the routing path between CD-4 and the CDM is complete. This allows the CD-4 to communicate indirectly with the CDM via CD-1.

3.2.2. Dynamic Reconfiguration. Figure 8(a) shows that if an intermediate node is released, any CDs indirectly connected to the CDM through that node cannot communicate with the CDM. Therefore, the middleware must be able to dynamically reconfigure the network topology. Figure 8(b) shows that CD-3, which has an error in one of its intermediate nodes, searches for neighboring CDs by broadcasting, and requests the routing table of CD-4. Figure 8(c) depicts CD-4 transmitting its routing table to CD-3. Figure 8(d) displays the new, completed path established by CD-3. In Figure 8, CD-3 appears to only receive CD-4's routing table, but it actually receives the routing table from each CD that received its broadcast search message. CD-3 selects the routing table of CD-4 because it provides the shortest routing path to connect with the CDM. A new routing table for CD-3 is generated as shown in Table 6.

3.3. Performance Evaluation. To test the EDCM proposed in this paper, we constructed the testbed environment illustrated in Figure 9. To represent the global network, two regional networks were constructed using three PC routers

with XORP [12] router software installed. We used four laptops, three netbooks, and two tablets as EDCM CDs.

This test bed assumes that Laptop 1 uses Bluetooth, Laptops 2, 3, and 4 use IPv6, Netbooks 1 and 3 use IPv6 and Bluetooth, Netbook 2 uses IPv6 and IPv4, Tablet 1 uses IPv4, and Tablet 2 uses Bluetooth. We confirmed that control messages are normally transmitted from Laptop 1 and Tablets 1 and 2, which are not IPv6-enabled devices, to Laptop 4, which belongs to a different local area network. Our experiment confirmed that when Netbook 1 or Netbook 3 was disconnected, the routing tables on Laptop 1 and Tablet 2, respectively, were reconfigured. This test demonstrates the stability of the proposed efficient heterogeneous network-routing scheme based on the dynamic control middleware.

4. Conclusions

Many industries rely on the ability to connect numerous devices to heterogeneous networks to deliver fundamental services, as observed in health care. However, DCM, a middleware for existing cyber-physical systems, only supports IPv6. Therefore, the many devices that do not use IPv6 cannot connect to a DCM-based CSP. In this paper, a network-routing module is designed to support CDM connections through IPv6 and other various network protocols such as IPv4, Bluetooth, and ZigBee. Through the proposed network-routing module, devices with heterogeneous protocols can be registered as components of an EDCM. Furthermore, if an intermediate node in a routing path is removed from the EDCM network, the network recovers dynamically.

Regarding future work, an EDCM network is constructed on a large scale, so it is necessary to study techniques for efficient traffic distribution. Also, authentication techniques to enable only authorized users to control the EDCM controller must be explored. Finally, we also plan to develop an encryption technique suitable for low-performance embedded devices.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean Government (Ministry of Science and ICT) (No. 2017R1C1B5075856).

References

[1] E. A. Lee, "Cyber physical systems: design challenges," Tech. Rep. UCB/EECS-2008-8, University of California, Berkeley, CA, USA, 2008.

[2] E. A. Lee, "Cyber-physical systems—are computing foundations adequate?," in *Position Paper for NSF Workshop On Cyber-Physical Systems: Research Motivation, Techniques and Roadmap*, pp. 1–9, Austin, TX, USA, October 2006.

[3] E. A. Lee, "Computing foundations and practice for cyber-physical systems: a preliminary report," Tech. Rep. UCB/EECS-2007-72, University of California, Berkeley, CA, USA, 2007.

[4] S. O. Park, T. H. Do, Y. S. Jeong, and S. J. Kim, "A dynamic control middleware for cyber physical systems on an IPv6-based global network," *International Journal of Communication Systems*, vol. 26, no. 6, pp. 690–704, 2013.

[5] S. W. Ahn, C. Yoo, S. H. Lee, H. S. Lee, and S. J. Kim, "Implementing virtual platform for global-scale cyber physical system networks," *International Journal of Communication Systems*, vol. 28, no. 13, pp. 1899–1920, 2015.

[6] T. Yang, Z. Huang, H. Pen, and Y. Zhang, "Optimal planning of communication system of CPS for distribution network," *Journal of Sensors*, vol. 2017, Article ID 9303989, 10 pages, 2017.

[7] N. Jabeur, N. Sahli, and S. Zeadally, "Enabling cyber physical systems with wireless sensor networking technologies, multiagent system paradigm, and natural ecosystems," *Journal of Sensors*, vol. 2015, Article ID 908315, 15 pages, 2015.

[8] M. C. Kim and S. J. Kim, "A scenario-based user-oriented integrated architecture for supporting interoperability among heterogeneous home network middlewares," in *Computational Science and Its Applications—ICCSA 2006. ICCSA 2006. Lecture Notes in Computer Science, vol 3983*, M. L. Gavrilova, O. Gervasi, V. Kumar, C. J. Kenneth, T. D. Taniar, A. Laganà, Y. Mun, and H. Choo, Eds., pp. 669–678, Springer, Berlin, Heidelberg, 2006.

[9] Y. W. Kim, H. S. Jang, J. H. Choi, C. H. Song, and Y. I. Eom, "Design of a multi-middleware bridge for supporting interoperability in home network environments," in *2008 International Conference on Advanced Language Processing and Web Information Technology*, pp. 477–482, Dalian Liaoning, China, July 2008.

[10] K.-D. Moon, Y.-H. Lee, C.-E. Lee, and Y.-S. Son, "Design of a universal middleware bridge for device interoperability in heterogeneous home network middleware," *IEEE Transactions on Consumer Electronics*, vol. 51, no. 1, pp. 314–318, 2005.

[11] S. O. Park, J. S. Kim, and S. J. Kim, "An object-based middleware supporting efficient interoperability on a smart home network," *Multimedia Tools and Applications*, vol. 63, no. 1, pp. 227–246, 2013.

[12] XORP, "XORP user manual version 1.6," http://www.xorp.org/.

Hindawi
Journal of Sensors
Volume 2018, Article ID 5948146, 8 pages
https://doi.org/10.1155/2018/5948146

Research Article

Intrusion Detection System Based on Evolving Rules for Wireless Sensor Networks

Nannan Lu,[1] **Yanjing Sun**,[1] **Hui Liu**,[2] **and Song Li**[1]

[1]*School of Information and Electrical Engineering, China University of Mining and Technology, Xuzhou, Jiangsu, China*
[2]*Institute of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, Beijing, China*

Correspondence should be addressed to Nannan Lu; lnn_921@126.com

Received 22 December 2017; Accepted 22 March 2018; Published 30 April 2018

Academic Editor: Mucheol Kim

Human care services, as one of the classical Internet of things applications, enable various kinds of things to connect with each other through wireless sensor networks (WSNs). Owing to the lack of physical defense devices, data exchanged through WSNs such as personal information is exposed to malicious attacks. Therefore, intrusion detection is urgently needed to actively defend against such attacks. Intrusion detection as a data mining procedure cannot control the size of rule sets and distinguish the similarity between normal and intrusion network behaviors. Therefore, in this paper, an evolving mechanism is introduced to extract the rules for intrusion detection. To extract diversified rules as well as control the quantity of rulesets, the extracted rules are examined according to the distance between the rules in the rule set of the same class and the rules in the rule set of different classes. Thereby, it alleviates the problem that the quantity of rules expands unexpectedly with the evolving genetic network programming. The simulations are conducted on a benchmark intrusion dataset, and the results show that the proposed method provides an effective solution to evolve the class association rules and improves the intrusion detection performance.

1. Introduction

The Internet of things (IoT) enables a large number of physical things or objects to connect, communicate, and exchange data with each other. IoT techniques span from health care to tactical military, in which human care is a type of classical application. The objects of human care services could include various kinds of medical equipment, even body parts. Wireless sensor networks (WSNs) are crucial for connecting, communicating, and exchanging data among such a large number of things. Although WSNs have the advantages of low installation cost, unattended network operation, and flexible deployment, their deficiency in physical defense devices renders both network and information vulnerable for malicious attacks [1]. To protect human care services from the internal or external attacks, prevention and detection are two main components involved in WSN security. However, as a passive network security mechanism, prevention is aimed at preventing any attack before it occurs and is therefore not sufficient. Thus, an active technique is urgently required to perceive malicious intrusions. Naturally, the focus shifts on the intrusion detection that can detect the actions attempting to compromise the confidentiality, integrity, or availability of one resource.

In general, the intrusion detection has two main techniques: misuse detection and anomaly detection. Misuse detection essentially identifies the previously known attacks from the normal network behaviors, while anomaly detection establishes the normal profiles to detect the new attacks. The combination of these two intrusion detection techniques is the hybrid intrusion detection. All the three techniques have been widely used in IoT. For example, Faisal et al. implemented anomaly detection to detect the external and internal attacks on smart meters [2]. Wang et al. and Pan et al. utilized the hybrid intrusion detection framework to protect the heterogeneous WSN, which was applied to power systems [3, 4]. Whereas, the specific methods of intrusion detection must be reviewed from the classical applications in the wired networks. Early studies on intrusion detection were conducted

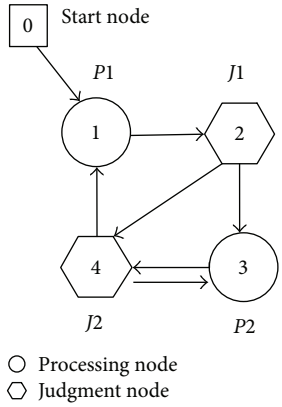


Figure 1: Phenotype of GNP.

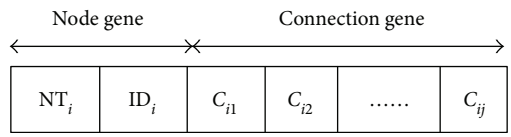


Figure 2: Genotype expression of GNP.

by Denning and Anderson [5, 6]. They aimed to build the monitoring systems for computer security, so that utilized statistics and rules to recognize attacks or viruses from the audited data. Since then, machine learning, data mining, statistic modeling, and pattern matching have been used to construct intrusion detection systems [7]. SVM was used to classify and select the audited data for the intrusion detection [8, 9]. The k-nearest neighbors (KNN) method has also been used to identify the intrusions through measuring distance [10]. Neural networks were also used to realize the intrusion detection systems, such as multilayer perceptron (MLP) [11]. Moreover, MLP was considered the basic unit to form the ensemble classifiers [12] such as AdaBoost. In [13], decision stumps were utilized as the weak classifiers to form a strong classifier.

Data mining is a successful solution to actively detect intrusive attacks based on the rules hidden in the network behavior data. Association rule mining is used to discover the correlations among the attribute sets in the data set for intrusion detection. The rules usually form as "$X \rightarrow Y$," which means that the -tuples in the dataset satisfy X is likely to satisfy Y. A RIPPER approach is proposed to generate frequent episodes firstly and then form the rules by associating the frequent episodes [14]. To extract the diversified rules, the fuzzy set theory was widely used to extract compact association rules. Tajbakhsh et al. [15] proposed a fuzzy association rule induction algorithm with two steps. The first step involved finding the significant itemsets with a higher significance factor than the user-specified threshold, and the second step involved generating rules by using the large itemsets induced in the first step. From the other perspective, intrusion detection generally distinguishes normal behavior, known intrusions and unknown intrusions, respectively, which can be taken as a classification procedure. Thus, the classes are considered with association rules

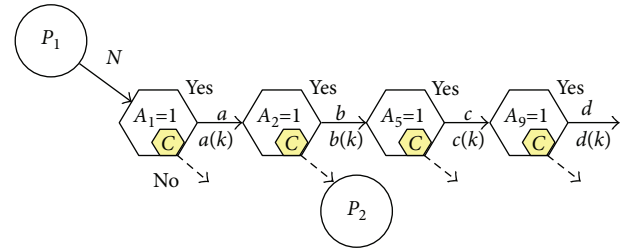


Figure 3: Class association rule mining based on GNP.

to form the class association rules. Different from the association rule, class association rule has the specified class label as its consequent part. Ozyer et al. [16] proposed to use GA boosting to find fuzzy class association rules. They encoded the rules as strings and used GA to evolve them. To extract as many rules as possible for identifying various kinds of intrusions, many algorithms were designed and implemented. Tsang et al. [17] employed a hierarchical GA structure. Each chromosome comprises control and parameter genes, and the parameters of fuzzy member functions were encoded as the parameter genes; the activations of which are managed by the control genes. Thus, the method was also used as a genetic feature selection wrapper to search for an optimal feature subset for dimensionality reduction. Feature selection can be used not only to alleviate the disadvantage of dimensionality and minimize the classification errors but also to improve the interpretability of the rule-based classifiers. Genetic programming (GP) has also been applied to intrusion detection. Some researches utilized GP to extract rules for intrusion detection based on its linear genomes and homologous crossover operators [18, 19]. In conclusion, most of current researches generally pursue the extraction of a large number of rules and overlook the discrimination of the rules [20]. Therefore, it is difficult to identify various types of intrusions with a high detection rate and a low false alarm rate. This could be due to the following two reasons. First, the network behavior data generated rapidly prompts the increase of the rules. Second, the similarity between the normal behavior and the new intrusion behavior limits the discrimination of the rules. Furthermore, this also brings about a considerable amount of redundant, irrelevant, and obvious information into the rule sets. In this case, the important rules are overwhelmed with the useless information. Therefore, the balance between the quality and quantity of rules is crucial for improving the intrusion detection performance.

To improve the quality of the rule sets as well as reserve the diversity of the rules, a new class association rule selection method is proposed based on genetic network programming (GNP) to solve the intrusion detection problem in smart human care services. Specifically, the similarities between rules and between rule sets are checked based on the distances during GNP evolution. The distance between the rules in the rule set of the same class is minimized, and the distance between the rules in the rule sets of the different classes is maximized by adding the newly extracted rules into the rule sets. If the above minimization and maximization criteria are

Table 1: Measurements of class association rules.

Class association rules	Support	Confidence
$(A_1 = 1) \Rightarrow (C = k)$	$a(k)/N$	$a(k)/a$
$(A_1 = 1) \wedge (A_2 = 1) \Rightarrow (C = k)$	$b(k)/N$	$b(k)/b$
$(A_1 = 1) \wedge (A_2 = 1) \wedge (A_5 = 1) \Rightarrow (C = k)$	$c(k)/N$	$c(k)/c$
$(A_1 = 1) \wedge (A_2 = 1) \wedge (A_5 = 1) \wedge (A_9 = 1) \Rightarrow (C = k)$	$d(k)/N$	$d(k)/d$

satisfied, the extracted rules are added to the rule sets; otherwise, they are discarded. In this way, redundant information can be avoided during the rule evolution, and the discrimination of the rule sets would be enhanced. In addition, this method also reserves the diversity of the rule sets according to the evolving mechanism. Thus, the GNP evolving method has the ability to discover discriminative class association rules for intrusion detection, which can further improve the intrusion detection performance.

The remainder of this paper is organized as follows. Section 2 describes the GNP structure and GNP-based class association rule mining in detail, and Section 3 introduces how to evolve class association rules based on distance. The simulation results are shown in Section 4, and Section 5 concludes this paper.

2. Genetic Network Programming

2.1. Basic Structure of GNP. GA has a string structure, and GP has a tree structure. With the complexity of problems increasing, it is difficult to express the problem using GA, and the GP structure starts bloating. As an extension of GA and GP, GNP has a quite different structure from GA and GP, which is the directed graph structure [21]. Figure 1 shows the phenotype of GNP, and there are three kinds of nodes in each individual. The start node is used to determine the first node to be executed. Judgment nodes work as the decision-making functions and are represented as $J_1, J_2, \ldots, J_m$. Processing nodes represent the functions of actions or processes and are expressed as $P_1, P_2, \ldots, P_n$. Node transition starts from the start node, and then, the next node to be executed is determined by the node transition. In addition, the number of nodes and their functions depend on the specific problem, which are determined by designers. In addition, judgment nodes have conditional branches, whereas processing nodes do not.

Figure 2 illustrates the genotype of the GNP structure.N T_i indicates the node type, the values of which are 0, 1, or 2. 0 is the start node, 1 is the processing node, and 2 is the judgment node. ID_i serves as an identification number. And C_{ij} denotes the node connection between node i and j.

2.2. Class Association Rule Mining Based on GNP. When GNP is used to extract class association rules, the function of the judgment node corresponds to the attribute of each tuple in the dataset, and the processing nodes are used to calculate the measurements of the class association rules. The specific procedure of class association rule mining using GNP is shown in Figure 3. GNP examines the attribute values of tuples in the dataset using judgment nodes and calculates the measurements using processing nodes. The judgment node determines the next node by the judgment result of yes or no, corresponding to the yes side or no side.

The yes side of the judgment node is connected to another judgment node. Judgment nodes can be reused and shared with other class association rules because of GNP's reusability. The no side of the judgment node is connected to another processing node, which represents the end of the current rule and the start of another new rule. The start node is connected to the first processing node. The connections of judgment nodes in Figure 3 are extracted as the candidate class association rules, which are shown below. There are four class association rules that correspond to four connections.

$$\begin{aligned} &(A_1 = 1) \Rightarrow (C = k), \\ &(A_1 = 1) \wedge (A_2 = 1) \Rightarrow (C = k), \\ &(A_1 = 1) \wedge (A_2 = 1) \wedge (A_5 = 1) \Rightarrow (C = k), \\ &(A_1 = 1) \wedge (A_2 = 1) \wedge (A_5 = 1) \wedge (A_9 = 1) \Rightarrow (C = k). \end{aligned} \tag{1}$$

To evaluate the above candidates of class association rules, we can calculate the corresponding support and confidence, which are shown in Table 1.

Let A_i be the item in the data set, and let its value be 1 or 0. Let C be the class label. So, the class association rule can be represented as the following unified form.

$$(A_p = 1) \wedge \cdots \wedge (A_q = 1) \Rightarrow (C = k), \quad k = 0, 1, 2, \ldots, K, \tag{2}$$

where $A_m = 1$ means that attribute A_m equals to 1 and C is the set of suffixes of classes.

3. Evolving Class Association Rules

GNP can extract a great number of class association rules for intrusion detection. However, with an increase in the amount of rules, the detection performance is not always enhanced by the extracted rules. Lots of rules bring redundant and irrelevant information into rule sets. This section describes how to implement the new evolving mechanism on the class association rule mining by GNP.

3.1. Jaccard Distance. Evolving class association rules are aimed at pruning the redundant and irrelevant rules for intrusion detection and at reserving the discriminative rules. In fact, a class association rule is composed of a set of attributes. Thus, the difference between two rules can be regarded as the distance between two sets of attributes, which is computed by the definition of Jaccard distance [22] shown as Definition 1.

```
Input: Target generation of GNP, N;
       Training data base, TrainDB;
Output: Accumulated ruleset, R;
for i = 0; i < N; i++ do
    while ith GNP generation do
        Extract next rule r based on TrainDB
        Distance_N ← CalculateDistance(r, R_N)
        // Calculate distance using the latest normal ruleset
        if Distance_N > CalculateDistance(r, R_N) then
            if the class of rule r equals normal then
                AddOneRule(r, R_N)
            else
                AddOneRule(r, R_I)
            end if
            break
        end if
        Distance_I ← CalculateDistance(r, R_I)
        // Calculate distance using the latest intrusion ruleset.
        if Distance_I > CalculateDistance(r, R_I) then
            if the class of rule r equals normal then
                AddOneRule(r, R_N)
            else
                AddOneRule(r, R_I)
            end if
        end if
    end while
    GNP population comets to (i + 1)th generation
end for
R ← MergeRulePool(R_N, R_I)
return R
```

Algorithm 1: (GNP with rule evolving).

Definition 1. Given two sets A and B, the Jaccard distance is defined as

$$D_J(A,B) = \frac{|A \cup B| - |A \cap B|}{|A \cup B|}, \tag{3}$$

where $|A \cup B|$ states the union of set A and set B, and $|A \cap B|$ indicates the intersection between set A and set B. The Jaccard distance can measure the degree of overlap between the two sets.

3.2. Rule Selection Based on Distance. Pruning the redundant and irrelevant rules is achieved by minimizing the distance of rules in the same class rule set as well as maximizing the distance between the different class rule sets. Therefore, the generated rules are checked according to the similarity of the rules and that of the rule sets. Specifically, when a newly extracted rule is added into the rule set, the distance between the rules in the rule set of the same class is minimized and the distance between the rules in the rule sets of different classes are simultaneously maximized. In this case, the rule is regarded as a distinguishable class association rule.

As the description of a class association rule, it comprises a group of attributes, which can be regarded as the mathematic theory "set." Thus, the distance either between the rules or between the rule sets can be described by the difference of two sets. Based on this principle, the distance between rule r and r' is defined as (4). And the distance between the rule sets can be calculated based on (4). The detailed definition is shown as (5).

$$d\left(r, r'\right) = \frac{|A_r \cup A_{r'}| - |A_r \cap A_{r'}| + \sum_{a \in A_r \cap A_{r'}} d\left(v(r,a), v\left(r', a\right)\right)}{|A_r \cup A_{r'}|}, \tag{4}$$

$$d\left(R, R'\right) = \frac{\sum_{r \in R} \sum_{r' \in R'} d\left(r, r'\right)}{|R|\left|R'\right|}, \tag{5}$$

where R and R' denote the rule set with different classes. r and r' represent the two rules. A_r and $A_{r'}$ are the corresponding attribute sets of rule r and rule r', respectively. a denotes the attribute in the rule. $v(r,a)$ is the value of attribute a of rule r. $d(R,R')$ stands for the distance between rule set R and rule set R'. $d(r,r')$ represents the distance between rule r and rule r', whose range is [0,1]. $d(v(r,a), v(r',a))$ is defined as

$$d\left(v(r,a), v\left(r', a\right)\right) = \begin{cases} 1, & v(r,a) \neq v\left(r', a\right), \\ 0, & v(r,a) = v\left(r', a\right). \end{cases} \tag{6}$$

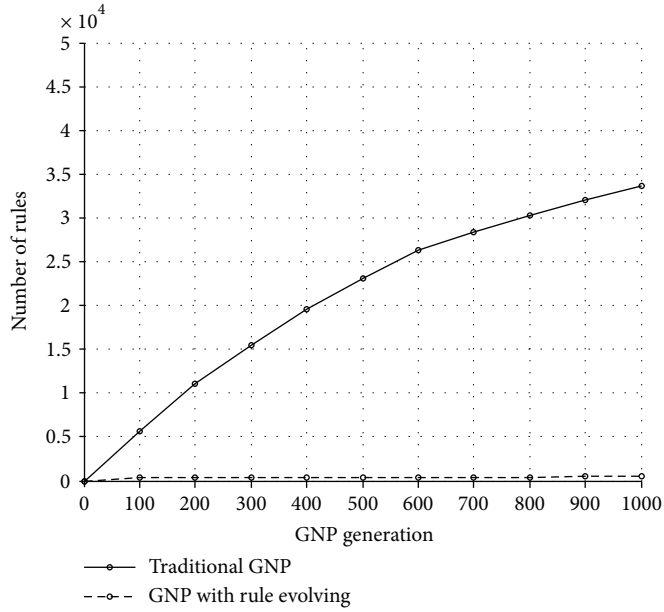


Figure 4: The comparison of rule quantity between the traditional GNP and GNP with rule evolving.

From (4) and (5), the modified distance considers the actual value of each attribute by adding $d(v(r,a), v(r',a))$ to the traditional Jaccard distance. $d(v(r,a), v(r',a)) = 0$ indicates that the attributes of rule r are completely the same with those of rule r', whereas $d(v(r,a), v(r',a)) = 1$ means that the attributes of rule r are completely different from those of rule r'. Therefore, the larger is the number of the same pairs (attribute, value), the shorter is the distance between r and r'. In this paper, the thresholds of intradistance between the rules in the same rule sets and interdistance between the rules in the different rule sets are all set as 0.98.

3.3. Evolving Class Association Rules. Except for support and confidence, χ^2 is also used to measure the significance of a rule. The class association rule is abbreviated as the form $X \rightarrow Y$, where $X, Y \subseteq I$ and $X \cap Y = \varnothing$, with I being the set of attributes. X and Y are the antecedent and consequent-of the association rule, respectively. Unlike the association rule, the class association rule has a class label as the consequent part. In this way, the support is defined as $\text{support}(X) = x$. x is the fraction of tuples containing X in the database. Confidence is defined by $\text{support}(X \cup Y)/\text{support}(X)$. Therefore, χ^2 of the rule is given by

$$\chi^2 = \frac{N(z - xy)^2}{xy(1-x)(1-y)}, \tag{7}$$

where N is the total number of tuples in the database, z is the value of $\text{support}(X \cup Y)$, and x and y are supports of X and Y, respectively.

Then, the minimum support, minimum confidence, and minimum χ^2 are used to select the candidate rules. After calculating the support, confidence, and χ^2 values of the above candidate class association rules, and if they satisfy the following conditions, $\text{support} \geq \text{support}_{\min}$, $\text{confidence} \geq \text{confidence}_{\min}$, and $\chi^2 \geq \chi^2_{\min}$, the rule is regarded as the important rule and then stored into the ruleset.

Table 2: Confusion matrix of classification results.

	Normal (C)	Misuse (C)	Anomaly (C)	Total
Normal (A)	8202	31	1478	9711
Misuse (A)	0	4776	546	5322
Anomaly (A)	1292	17	6202	7511
Total	9494	4824	8226	22,544

Each individual is evaluated by the fitness function defined by

$$\text{Fitness} = \sum_{r \in R} \left\{ \chi^2(r) + 10(n_{\text{ante}}(r) - 1) + \alpha_{\text{new}}(r) \right\}, \tag{8}$$

where R in $r \in R$ is the set of suffixes of the extracted important rules from the individuals, $n_{\text{ante}}(r)$ is the number of attributes in the antecedent part of rule r, and $\alpha_{\text{new}}(r)$ is the additional constant shown as (9)

$$\alpha_{\text{new}}(r) = \begin{cases} \alpha_{\text{new}}, & \text{when rule } r \text{ is newly extracted,} \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

Therefore, the fitness function of GNP is concerned with importance, complexity, and novelty of rule r.

The pseudocode of evolving class association rules by GNP is summarized in Algorithm 1.

4. Simulations

4.1. Data Set. Owing to the lack of realistic data sets of smart human care services, the benchmark data set NSL-KDD is used to verify the validity of the proposed method

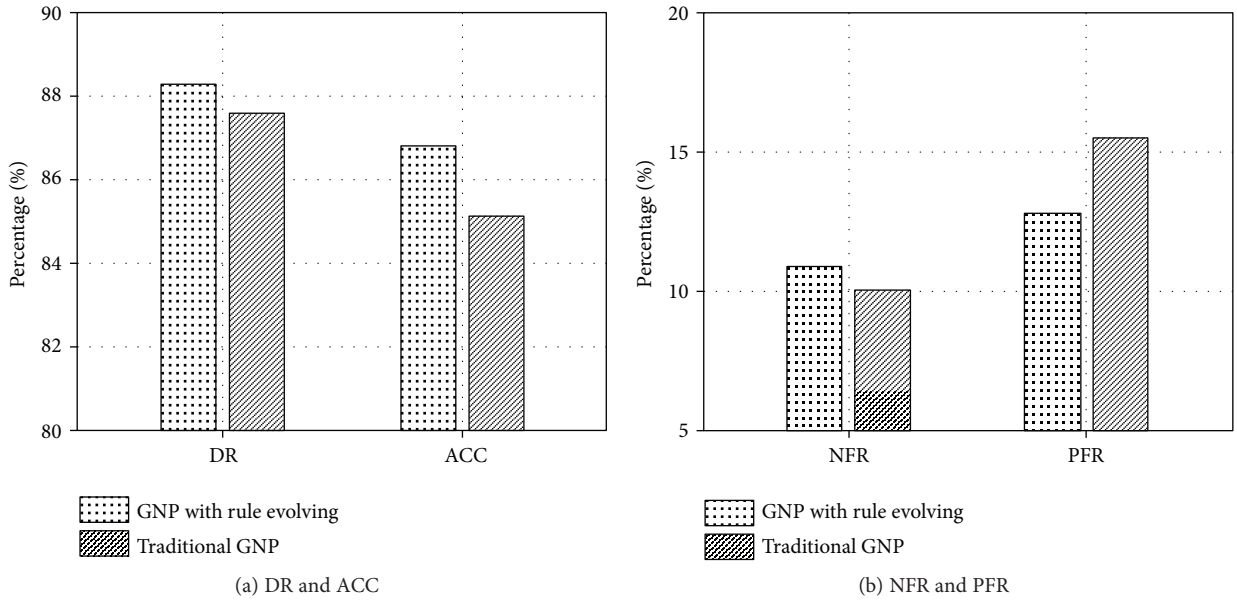


Figure 5: Detection performance comparisons of the traditional GNP and GNP with rule evolving.

Table 3: The accuracy comparisons of the traditional GNP and GNP with rule evolving (%).

Classifier	Traditional GNP (%)	GNP with rule evolving (%)
SVM	72.89	75.98
BP	71.87	74.82
MLP	73.53	74.91
KNN	75.69	77.30
Logic regression	73.56	74.83
Decision tree	79.63	80.22
AdaBoost	75.23	73.50
Naive Bayes	76.31	74.72
Cluster Gaussian	85.07	86.78
Average	**75.97**	**77.00**

[23, 24]. NSL-KDD is a new version of KDD CUP 1999 data set [25]. Both NSL-KDD and KDD CUP 1999 include a wide variety of intrusions simulated in a military network environment, which is difficult for a self-build simulation environment to acquire such diversified categories of intrusions. However, the NSL-KDD data set is different from KDD CUP 1999, which is composed of the most difficult detected data evaluated by the classical classification methods. Thus, NSL-KDD is a challenge data set for evaluating the intrusion detection methods. Moreover, compared to the KDD CUP 1999 data set, the intrusion detection performance on the NSL-KDD data set will not be biased towards the intrusions easily detected.

Each audit data in NSL-KDD consists of 41 attributes including continuous and discrete ones and one class label. Except for normal audit data, there are four types of attacks in this dataset, which are denial of service (DOS), probe, user to root (U2R), and root to local (R2L).

4.2. Parameter Settings. We use $\text{support}_{\min}(N) = 0.1$, $\text{support}_{\min}(I) = 0.075$, $\text{confidence}_{\min} = 0.8$, and $\chi^2_{\min} = 6.64$, where N and I indicate normal and intrusion, respectively. Class association rules are extracted for each class using GNP. The population size of GNP is 120. The number of processing nodes and judgment nodes are 10 and 100, respectively. In addition, the crossover rate is 1/5. The mutation rate for P_{m1} is 1/3 and for P_{m2} is 1/3, in which P_{m1} and P_{m2} mutate the connections of the branches and the contents of the nodes, respectively. The condition of termination is 1000 generations.

4.3. Result Analysis. First, we randomly select 2000 normal data and 2000 intrusion data as the training set. The testing set consists of 9711 normal data and 12,833 intrusion data. Both the training set and the testing set are from NSL-KDD, which avoids redundant records and improves the difficulty level of KDD Cup 1999. GNP-based class association rule mining is conducted on the prepared training set. The proposed method is compared with the traditional GNP-based class association rule mining shown in Figure 4. Different from GNP with rule evolving, the traditional GNP has no action of automatic selection of useful rules, which always extracts a great number of class association rules. After 1000 generations, the traditional GNP increasingly generates rules, while the proposed method has been converged already. We can conclude that GNP with rule evolving has the strong ability to reduce the rule quantity in rule mining. It can be also regarded as an online rule pruning scheme.

Furthermore, the detection performance of the proposed method on the NSL-KDD data set is investigated. Here, we use the classifier of cluster Gaussian referred to the literature [26]. The cluster Gaussian classifier utilizes

Table 4: The accuracy comparisons of CBA, CMAR, and GNP with rule evolving (%).

Classifier	Decision tree	SVM	KNN	AdaBoost	Cluster Gaussian
CBA	70.64	74.47	71.75	72.83	86.09
CMAR	72.39	72.28	72.99	72.19	85.48
GNP with rule evolving	80.22	75.98	77.30	73.50	86.78

the information of known normal and intrusion data to find the extract boundary of normal and known intrusions. In terms of the data distribution, it clusters the similar data which are supposed to have the similar network behaviors. And the classifier further uses Gaussian functions to find the cluster boundaries and data distribution to determine the cluster number. Table 2 shows the confusion matrix of the detection results. "A" is the actual class of the data and "C" is labeled by the classifier. From the confusion matrix, detection rate (DR), accuracy, positive false rate (PFR), and negative false rate (NFR) are calculated. DR indicates the rate of the data that are correctly classified into normal or intrusion. ACC (accuracy) is the rate of the data that are accurately classified as normal, misuse intrusion, or anomaly intrusion. PFR represents that the classifier identifies the normal data as misuse intrusion or anomaly intrusion. NFR represents that misuse and anomaly intrusions are identified as normal.

Therefore, according to Table 2, DR, accuracy, PFR, and NFR are calculated as follows.

$$\begin{aligned} \text{DR} &= \frac{8202 + 4776 + 6202 + 17 + 546}{22544} = 87.58\%, \\ \text{Accuracy} &= \frac{8202 + 4776 + 6202}{22544} = 85.08\%, \\ \text{PFR} &= \frac{31 + 1478}{9711} = 15.54\%, \\ \text{NFR} &= \frac{0 + 1292}{5322 + 7511} = 10.07\%. \end{aligned} \tag{10}$$

From the results, misuse intrusion and normal are easy to distinguish by the evolved rules. Though most of anomaly intrusions have been identified, a lot of anomaly intrusions are still difficult to detect. Furthermore, we compare the traditional GNP with the proposed method on DR, accuracy, NFR, and PFR. As shown in Figure 5, GNP with rule evolving obtains higher DR and ACC and lower PFR, but NFR is a little bit higher in GNP with rule evolving than in the traditional GNP. Therefore, anomaly intrusions are still difficult to distinguish. The similarities between anomaly intrusions and normal patterns account for this phenomenon.

In order to further demonstrate the proposed method, classical machine learning algorithms are taken as comparative classifiers, including support vector machine (SVM), back propagation (BP) neural network, multilayer perception (MLP), *k*-nearest neighbor (KNN), logic regression, decision tree, AdaBoost, naive Bayes, and cluster Gaussian.

Table 3 shows the detection accuracy of the traditional GNP and GNP with rule evolving based on different classifiers. By evolving 1000 generations, GNP extracts 33,723 rules and the proposed GNP extracts 436 rules. Then, 9 classifiers are used to evaluate the intrusion detection performance based on the two GNP. Among them, 7 classifiers on GNP with rule evolving acquire higher accuracy than those on the traditional GNP. In addition, the average accuracy of GNP with rule evolving is also better than that of the traditional GNP. The results demonstrate that the proposed method can evolve better rules for intrusion detection. Besides, we evaluate the proposed method by comparing it with the classical classification rule mining methods such as classification based on associations (CBA) [27] and classification based on multiple association rules (CMAR) [28]. Both CBA and CMAR contain the rule pruning procedure. With the default classifiers, CBA and CMAR obtain accuracies of 74.63% and 72.17%, respectively, which are lower than the average accuracy of 77% obtained using GNP with a rule selection mechanism. In addition, we select some of the classical classification methods to evaluate the effectiveness of the proposed method, which are the decision tree, SVM, KNN, AdaBoost, and cluster Gaussian. Table 4 illustrates the accuracy comparisons of CBA, CMAR, and GNP with rule evolving. As shown in Table 4, GNP with rule evolving has higher classification accuracies than the other rule-based methods. Thus, it is necessary to consider the rule evolving technique in the rule mining. And the rule evolving is capable of selecting useful rules and reducing the redundant and irrelevant rules.

5. Conclusions

In this paper, an intrusion detection system based on evolving class association rules is proposed as a security solution for smart human care services. In general, it utilizes a class association rule evolving strategy to construct the intrusion detection system. As a data mining solution, GNP with rule evolving can generate diversified class association rules and control the quantity of the rules simultaneously. For intrusion detection, the significance test is performed to ensure the importance of generated rules. In order to generate the more discriminative class association rules, the Jaccard distance is modified to measure the similarity between rules and different rule sets. In this way, the distance of the rules in the rule set with the same class is minimized and the distance between rules in the rule sets with different classes is maximized. The simulations conducted on the NSL-KDD dataset theoretically verify that GNP with rule evolving efficiently controls the quantity of generated rules and improves the detection performance by reducing the redundancy of the rules. In the future, we plan to verify the effectiveness of the

intrusion detection system on the self-build simulation environment.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by Basic Scientific Research Fund for Central Universities under Grant no. 2015QNB20, in part by the Natural Science Foundation of Jiangsu Province under Grant no. BK20150204, in part by China Postdoctoral Science Foundation under Grant no. 2015M581884, and in part by National Natural Science Foundation of China under Grant no. 51734009 and no. 51504255.

References

[1] I. Butun, S. D. Morgera, and R. Sankar, "A survey of intrusion detection systems in wireless sensor networks," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 266–282, 2014.

[2] M. A. Faisal, Z. Aung, J. R. Williams, and A. Sanchez, "Data-stream-based intrusion detection system for advanced metering infrastructure in smart grid: a feasibility study," *IEEE Systems Journal*, vol. 9, no. 1, pp. 31–44, 2015.

[3] S. S. Wang, K. Q. Yan, S. C. Wang, and C. W. Liu, "An integrated intrusion detection system for cluster-based wireless sensor networks," *Expert Systems with Applications*, vol. 38, no. 12, pp. 15234–15243, 2011.

[4] S. Pan, T. Morris, and U. Adhikari, "Developing a hybrid intrusion detection system using data mining for power systems," *IEEE Transactions on Smart Grid*, vol. 6, no. 6, pp. 3104–3113, 2015.

[5] D. E. Denning, "An intrusion-detection model," *IEEE Transactions on Software Engineering*, vol. SE-13, no. 2, pp. 222–232, 1987.

[6] J. P. Anderson, "Computer security threat monitoring and surveillance," Tech. Rep., James P. Anderson Co., Fort, Washington, PA, USA, 1980.

[7] S. X. Wu and W. Banzhaf, "The use of computational intelligence in intrusion detection systems: a review," *Applied Soft Computing*, vol. 10, no. 1, pp. 1–35, 2010.

[8] A. A. Aburomman and M. B. I. Reaz, "A novel weighted support vector machines multiclass classifier based on differential evolution for intrusion detection systems," *Information Sciences*, vol. 414, pp. 225–246, 2017.

[9] E. Kabir, J. Hu, H. Wang, and G. Zhuo, "A novel statistical technique for intrusion detection systems," *Future Generation Computer Systems*, vol. 79, pp. 303–318, 2018.

[10] C. Guo, Y. Ping, N. Liu, and S. S. Luo, "A two-level hybrid approach for intrusion detection," *Neurocomputing*, vol. 214, pp. 391–400, 2016.

[11] J. Cannady, "Artificial neural networks for misuse detection," in *Proceedings of the 1998 National Information Systems Security Conference*, pp. 443–456, Arlington, VA, USA, 1998.

[12] D. Parikh and T. Chen, "Data fusion and cost minimization for intrusion detection," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 3, pp. 381–389, 2008.

[13] W. Hu, W. Hu, and S. Maybank, "Adaboost-based algorithm for network intrusion detection," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 38, no. 2, pp. 577–583, 2008.

[14] W. W. Cohen, "Fast effective rule induction," in *Proceedings of the Twelfth International Conference on Machine learning*, pp. 115–123, Tahoe City, California, 1995.

[15] A. Tajbakhsh, M. Rahmati, and A. Mirzaei, "Intrusion detection using fuzzy association rules," *Applied Soft Computing*, vol. 9, no. 2, pp. 462–469, 2009.

[16] T. Ozyer, R. Alhajj, and K. Barker, "Intrusion detection by integrating boosting genetic fuzzy classifier and data mining criteria for rule pre-screening," *Journal of Network and Computer Applications*, vol. 30, no. 1, pp. 99–113, 2007.

[17] C. H. Tsang, S. Kwong, and H. Wang, "Genetic-fuzzy rule mining approach and evaluation of feature selection techniques for anomaly intrusion detection," *Pattern Recognition*, vol. 40, no. 9, pp. 2373–2391, 2007.

[18] J. V. Hansen, P. B. Lowry, R. D. Meservy, and D. M. Mcdonald, "Genetic programming for prevention of cyberterrorism through dynamic and evolving intrusion detection," *Decision Support Systems*, vol. 43, no. 4, pp. 1362–1374, 2007.

[19] D. Song, M. I. Heywood, and A. N. Zincir-Heywood, "Training genetic programming on half a million patterns: an example from anomaly detection," *IEEE Transactions on Evolutionary Computation*, vol. 9, no. 3, pp. 225–239, 2005.

[20] E. Lazcorreta, F. Botella, and A. Fernández-Caballero, "Towards personalized recommendation by two-step modified apriori data mining algorithm," *Expert Systems with Applications*, vol. 35, no. 3, pp. 1422–1429, 2008.

[21] S. Mabu, K. Hirasawa, and J. Hu, "A graph-based evolutionary algorithm: genetic network programming (GNP) and its extension using reinforcement learning," *Evolutionary Computation*, vol. 15, no. 3, pp. 369–398, 2007.

[22] M. Levandowsky and D. Winter, "Distance between sets," *Nature*, vol. 234, no. 5323, pp. 34-35, 1971.

[23] A. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A detailed analysis of the kdd cup 99 data set," in *2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications*, pp. 53–58, Ottawa, ON, Canada, 2009.

[24] http://nsl.cs.unb.ca/NSL-KDD/.

[25] http://kdd.ics.uci.edu/databases/kddcup99/.

[26] N. Lu, S. Mabu, Y. Li, and K. Hirasawa, "Classification with clustering and Gaussian functions in intrusion detection system," *IEEJ Transactions on Electronics Information and Systems*, vol. 134, no. 12, pp. 1908–1915, 2014.

[27] B. Liu, W. Hsu, and Y. Ma, "Mining association rules with multiple minimum supports," in *KDD '99 Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 337–341, San Diego, California, USA, 1999.

[28] W. Li, J. Han, and J. Pei, "Cmar: accurate and efficient classification based on multiple class-association rules," in *Proceedings 2001 IEEE International Conference on Data Mining*, pp. 369–376, San Jose, CA, USA, 2001.

Hindawi
Journal of Sensors
Volume 2018, Article ID 7501914, 8 pages
https://doi.org/10.1155/2018/7501914

Research Article

Multistandard Receiver Design for Telemedicine Monitoring System

Hongmei Wang[1], **Chong Yao**[1], **Faguang Wang**[1,2], **Shiyin Li**[1], **and Sanghyuk Lee**[3,4]

[1]*School of Information and Control Engineering, China University of Mining and Technology, Xuzhou, China*
[2]*Collaborative Innovation Center of Intelligent Mining Equipment, China University of Mining and Technology, No. 1 Daxue Road, Xuzhou 221116, China*
[3]*Electrical and Electronic Engineering, Xi'an Jiaotong-Liverpool University, Suzhou, China*
[4]*Ton Duc Thang University, Go Chi Minh, Vietnam*

Correspondence should be addressed to Faguang Wang; wfaguang@cumt.edu.cn

Received 23 December 2017; Accepted 26 March 2018; Published 30 April 2018

Academic Editor: Mucheol Kim

In short-distance wireless communications for telemedicine monitoring, different medical data measurement equipment has different wireless transmission modes. A multistandard receiver is designed that can adapt to different medical data measuring equipment. Using a second-order bandpass sampling for the design of antialiasing filters, two aliasing signals can be separated. Simultaneously, constraint conditions for sampling frequency are not as critical. The design is useful for a multistandard receiver in a telemedicine monitoring system and has the advantages such as saving spectrum resources and facilitating spectrum planning.

1. Introduction

With the rapid development of Internet of things technology and short-distance wireless communications technology, telemedicine monitoring network technology has become a hot research topic [1]. This technology allows doctors to remotely diagnose the condition of patients and provide timely medical advice. In the remote medical monitoring network, the transmission of local medical data, such as heartbeat, respiration, blood oxygen, and pulse, relies mainly on short-distance wireless communications technology such as Bluetooth, Wi-Fi, and ZigBee [2] These short-distance wireless communications technologies have different applications and transmission rates. They have their own advantages and disadvantages, and each technology is constantly adapting to different types or standards of medical data transmission. The objective of our study is to design a multistandard receiver that can adapt to different communication standards so that various medical data can be flexibly received.

At present, short-distance wireless communications technologies for telemedicine monitoring, such as Bluetooth, Wi-Fi, and ZigBee, all use the 2.4 GHz frequency range; therefore, spectrum resources are absent. Improving spectrum utilization is a very important problem in the receiver. To avoid aliasing caused by bandpass sampling (BPS), most researchers consider choosing the lowest possible sampling frequency in order to reduce the burden of subsequent digital processing without aliasing in the spectrum [3–5]. Many researchers have tried to find a new algorithm to simplify the frequency selection process [6–8]. However, preventing aliasing will reduce spectrum utilization and the complexity of the computation process will increase the difficulty of implementation.

This work proposes a solution for aliasing in receivers that can reduce the limitations in sampling frequency, improve spectrum utilization, and realize multistandard receivers. In previous work, a sampling frequency that is twice the signal bandwidth was used to receive two-band signals [9]. This method allowed two signals to overlap with each other and separate using interplants, which eases the restriction on sampling frequency. However, this method did not provide the constraint conditions for

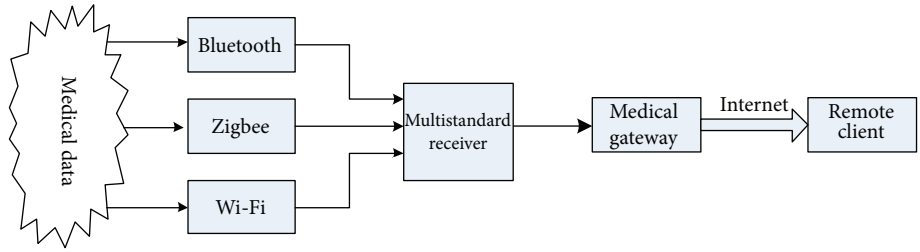


Figure 1: Structure of telemedicine monitoring system.

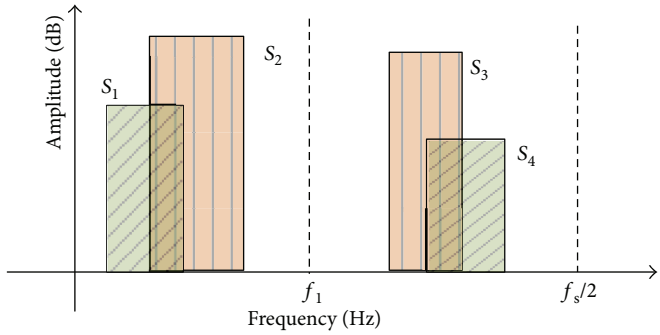


Figure 2: Spectra after BPS.

sampling frequency and the solution for aliasing by more than two signals. Based on this method, this paper gives the constraint conditions for frequency selection and a method to process aliasing by more than two standard signals.

This paper is organized as follows. Section 2 gives the structure of telemedicine monitoring system and the method to design antialiasing filters to suppress aliasing. In Section 3, aliasing analysis and constraint conditions are given and analyzed. Section 4 presents the simulation results for the multistandard receiver and spectral analysis. Section 5 tests the proposed algorithm in hardware. Section 6 provides concluding remarks.

2. Multistandard Receiver

2.1. Structure of Telemedicine Monitoring System. The structure of telemedicine monitoring system is shown in Figure 1.

Medical data, such as heartbeat, respiration, blood oxygen, and pulse, are measured by different equipment and transmitted by different wireless communications technology. The multistandard receiver samples the signals and removes aliasing. The received data are sent to the medical gateway and transmitted to the remote client via the Internet.

The core of the system is the design of an antialiasing module that can separate the overlapped signals.

2.2. Antialiasing Filter Design. Assume that the primary signal is sampled at a sampling rate f_s. Any signal in the frequency zone of index n expressed as [9]

$$\left(n-\frac{1}{2}\right)f_s < |f| < \left(n+\frac{1}{2}\right)f_s \tag{1}$$

is aliased into the first Nyquist zone $|f| < f_s/2$, which is the frequency zone with index zero.

Assume that four different standard signals, S_1', S_2', S_3', and S_4', with frequency zone index n_1, n_2, n_3, and n_4, respectively, are bandpass sampled at sampling frequency f_s simultaneously. The sampled spectra are represented as S_1, S_2, S_3, and S_4, respectively, as shown in Figure 2, where signals S_1 and S_2 overlap in area 0 to f_1, and signals S_3 and S_4 overlap in area f_1 to $f_s/2$.

In order to separate the overlapping signals, a second-order BPS sampling structure is designed, as shown in Figure 3. Two analog-to-digital converters (ADCs), ADC A and ADC B, are used for BPS, where ADC B introduces time delay T_Δ. The two sampled channels are referred to as channel A and channel B.

According to the characteristics of second-order BPS [4], the sampled signal spectra in the two channels satisfy the following equation:

$$S_{i\mathrm{B}} = S_{i\mathrm{A}} \cdot \beta^{-ni}, \tag{2}$$

where $S_{i\mathrm{A}}$ ($i = 1, 2, 3, 4$) represents the signals sampled in channel A, and $S_{i\mathrm{B}}$ represents the signals sampled in channel B, $\beta = e^{-j2\pi T_\Delta f_s}$. Because the negative and positive spectra are symmetrical, here, we discuss only the positive spectra.

In Figure 3, antialiasing filters F_{A1} and F_{B1} should be designed to suppress signals S_2 and S_4, and F_{A2} and F_{B2} should suppress signals S_1 and S_3.

In area 0 to f_1, S_2 needs to be suppressed. Thus, in the frequency domain, F_{A1} and F_{B1} should satisfy

$$\begin{aligned} S_{1\mathrm{A}}F_{\mathrm{A1}} + S_{1\mathrm{B}}F_{\mathrm{B1}} &= S_{1\mathrm{A}}, \\ S_{2\mathrm{A}}F_{\mathrm{A1}} + S_{2\mathrm{B}}F_{\mathrm{B1}} &= 0. \end{aligned} \tag{3}$$

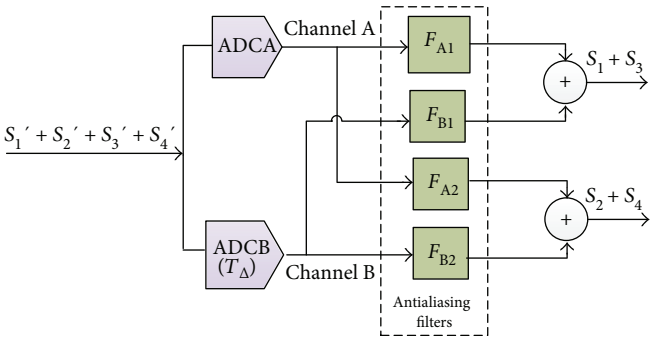


FIGURE 3: Structure of the second-order BPS system.

After simplification,

$$\begin{aligned} F_{A1} + \beta^{-n1} F_{B1} &= 1, \\ F_{A1} + \beta^{-n2} F_{B1} &= 0. \end{aligned} \tag{4}$$

From (2) and (3), we get

$$\begin{aligned} F_{A1} &= \frac{-\beta^{-n2}}{\beta^{-n1} - \beta^{-n2}}, \\ F_{B1} &= \frac{1}{\beta^{-n1} - \beta^{-n2}}, \\ f &\in [0, f_1]. \end{aligned} \tag{5}$$

In area f_1 to $f_s/2$, S_4 needs to be suppressed. Thus, in the frequency domain, F_{A1} and F_{B1} should satisfy

$$\begin{aligned} S_{3A} F_{A1} + S_{3B} F_{B1} &= S_{3A}, \\ S_{4A} F_{A1} + S_{4B} F_{B1} &= 0. \end{aligned} \tag{6}$$

After simplification,

$$\begin{aligned} F_{A1} + \beta^{-n3} F_{B1} &= 1, \\ F_{A1} + \beta^{-n4} F_{B1} &= 0. \end{aligned} \tag{7}$$

From (2) and (7), we get

$$\begin{aligned} F_{A1} &= \frac{-\beta^{-n4}}{\beta^{-n3} - \beta^{-n4}}, \\ F_{B1} &= \frac{1}{\beta^{-n3} - \beta^{-n4}}, \\ f &\in \left[f_1, \frac{f_s}{2}\right]. \end{aligned} \tag{8}$$

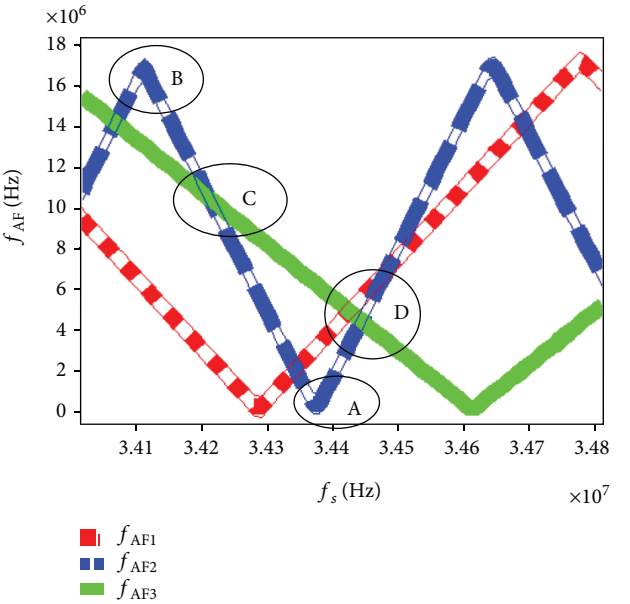


FIGURE 4: The relationship between signal mirroring frequency and sampling frequency.

Using the same method, we can get the expressions for F_{A2} and F_{B2} in the frequency domain, as shown as follows:

$$\begin{aligned} F_{A2} &= \frac{-\beta^{-n1}}{\beta^{-n2} - \beta^{-n1}}, \\ F_{B2} &= \frac{1}{\beta^{-n2} - \beta^{-n1}}, \\ f &\in [0, f_1], \\ F_{A2} &= \frac{-\beta^{-n3}}{\beta^{-n4} - \beta^{-n3}}, \\ F_{B2} &= \frac{1}{\beta^{-n4} - \beta^{-n3}}, \\ f &\in \left[f_1, \frac{f_s}{2}\right]. \end{aligned} \tag{9}$$

3. Constraint Conditions

3.1. Aliasing Analysis. Assume an RF signal with central frequency f_c bandpass sampled at sampling frequency f_s, the central frequency in the first Nyquist zone (called mirroring frequency) can be defined as

$$f_{AF} = \begin{cases} f_c - \dfrac{r f_s}{2}, & r = \left\lfloor \dfrac{f_c}{f_s/2} \right\rfloor \text{ is even}, \\ \dfrac{(r+1) f_s}{2} - f_c, & r = \left\lfloor \dfrac{f_c}{f_s/2} \right\rfloor \text{ is odd}. \end{cases} \tag{10}$$

where $\lfloor \bullet \rfloor$ means to round down.

Assume three bandpass signals that are received simultaneously with mirroring frequency f_{AF1}, f_{AF2}, and f_{AF3}.

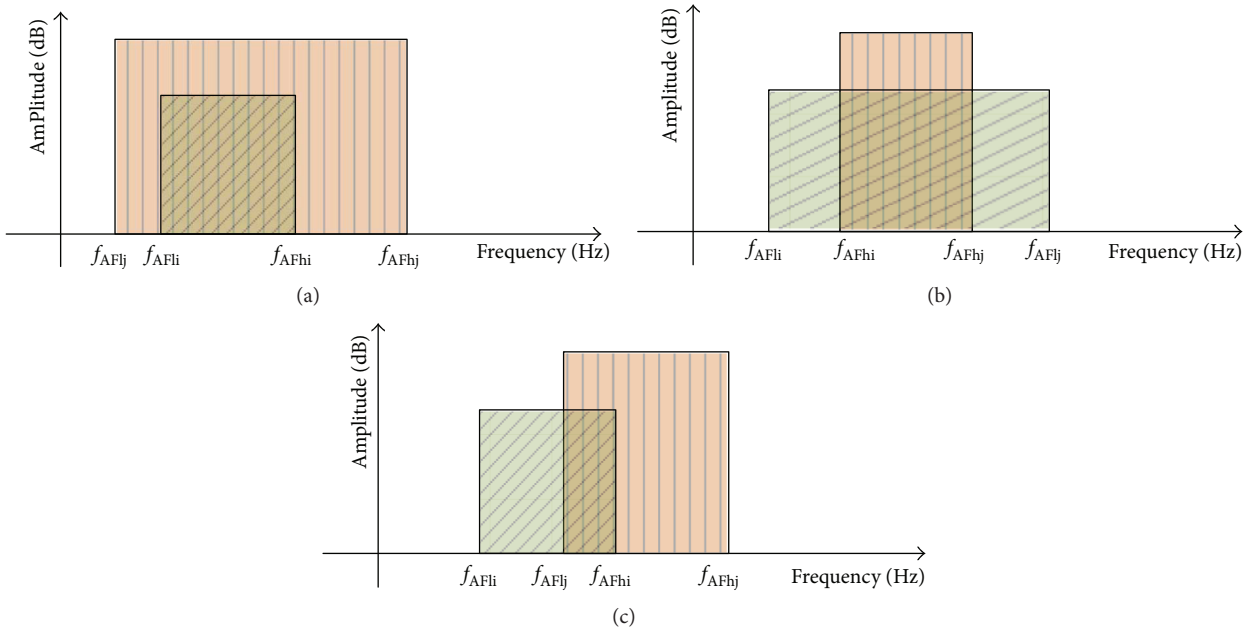


Figure 5: (a) Aliasing diagram (Case a). (b) Aliasing diagram (Case b). (c) Aliasing diagram (Case c).

Table 1: Parameters of four input signals.

Number	Central frequency	Signal types	Frequency zone index	Frequency offset
S_1	2.428 GHz	Bluetooth	101	4 MHz
S_2	2.452 GHz	Wi-Fi	102	4 MHz
S_3	2.408 GHz	ZigBee	100	8 MHz
S_4	2.464 GHz	Bluetooth	103	8 MHz

Figure 4 shows the relationship between the mirroring frequency and sampling frequency f_s.

Consider signal f_{AF2}, with different selections of f_s, it has four kinds of aliasing:

(i) Aliasing by self-image spectrum around zero frequency, as shown in area A of Figure 4

(ii) Aliasing by self-image spectrum around $f_s/2$ frequency, as shown in area B of Figure 4

(iii) Aliasing by another signal, as shown in area C of Figure 4

(iv) Aliasing by another two signals, as shown in area D of Figure 4

The traditional way of selecting f_s is to avoid all kinds of aliasing, which limits the area of f_s, and usually needs a large amount of calculations. In Section 2, aliasing caused by image spectrum and overlap with two signals can be solved by software. Hence, when we select f_s, less limiting conditions are allowed.

3.2. Constraint Condition. According to the abovementioned rules, the sampling rate constraint conditions are given to allow the presence of two aliasing signals at the same location after sampling. The details can be expressed by the following equations:

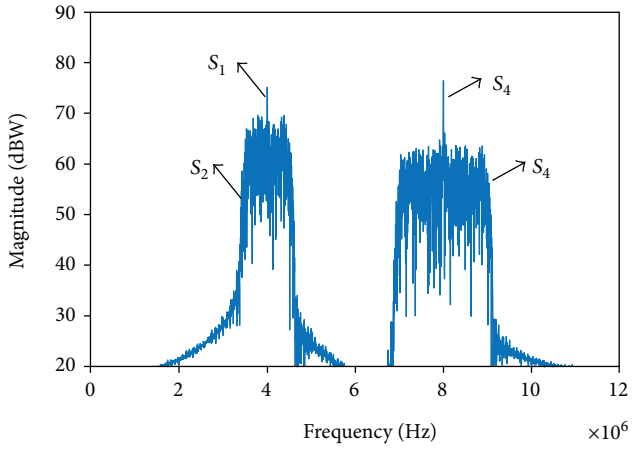


Figure 6: Signal spectra after BPS.

$$f_s \geq 2[B_{\max} + 2B_G]. \tag{11}$$

If

$$f_{DFi} < \frac{B}{2 + B_G}, \tag{12}$$

then

$$\frac{B_j}{2 + B_G} < f_{DFj.} \tag{13}$$

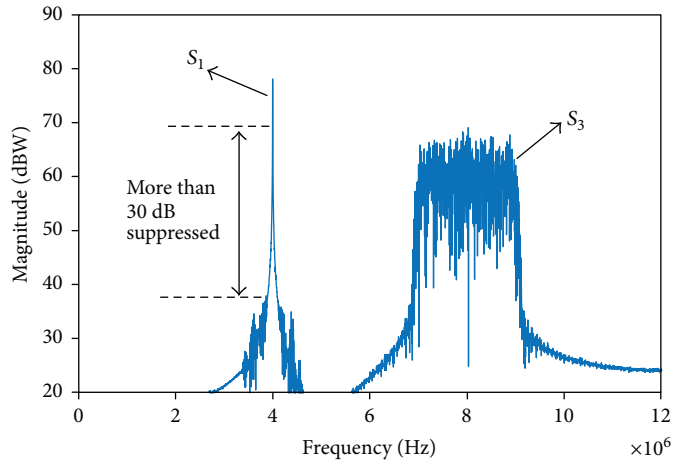


FIGURE 7: Signal spectra after antialiasing filters F_{A1} and F_{B1}.

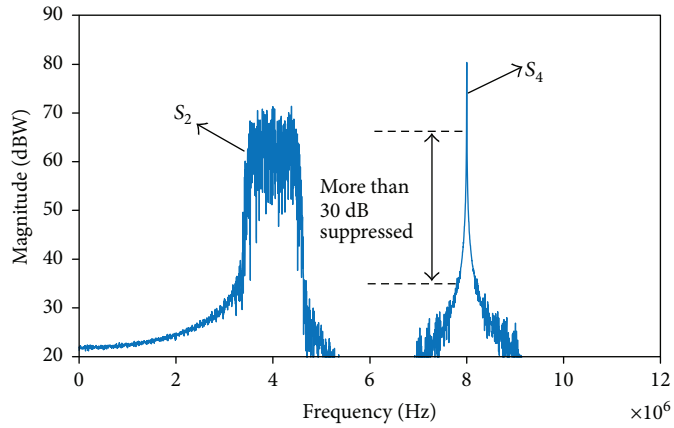


FIGURE 8: Signal spectra after antialiasing filters F_{A2} and F_{B2}.

If

$$f_{DFi} > \frac{f_s}{2} - \left(\frac{B_i}{2 + B_G}\right), \tag{14}$$

then

$$f_{DFj} < \frac{f_s}{2} - \left(\frac{B_j}{2 + B_G}\right), \tag{15}$$

where $i = 2, 3, \ldots, N$ and $j = 2, 3, \ldots, i - 1$.

If $f_{AFli} > f_{AFlj}$, then $f_{AFhk} < f_{AFli}$ or $f_{AFlk} > f_{AFhi}$; if $f_{AFli} < f_{AFlj}$ and $f_{AFhi} > f_{AFhj}$, then $f_{AFhk} < f_{AFlj}$ or $f_{AFlk} > f_{AFhj}$; if $f_{AFli} < f_{AFlj}$ and $f_{AFlj} < f_{AFhi} < f_{AFhj}$, then $f_{AFlk} < f_{AFlj}$ or $f_{AFlk} > f_{AFhi}$, where $i = 2, 3, \ldots, N$, $j = 1, 2, 3, \ldots, i - 1$, $k \in [1, N]$, $k \neq i$, and $k \neq j$.

In the above constraint conditions, B_{max} represents the maximum bandwidth of N bandpass signals; B_G represents protection bandwidth; B_i, f_{DFi}, f_{AFli}, and f_{AFhi} represent the bandwidth, central frequency, minimum frequency, and maximum frequency of signal i, respectively, in the first Nyquist region after BPS.

Constraint 1. It indicates that the sampling frequency needs to be two times greater than the maximum bandwidth of the signal (plus protection bandwidth).

Constraint 2. It means that only one signal is allowed to have zero-bound aliasing or $f_s/2$-bound aliasing, that is, the spectral aliasing of the image itself.

Constraint 3. It means that only two aliasing signals are allowed in the same location. There are three cases for this constraint. Case a $f_{AFli} > f_{AFlj}$ is shown in Figure 5(a). Case b $f_{AFli} < f_{AFlj}$ and $f_{AFhi} > f_{AFhj}$ is shown in Figure 5(b). Case c $f_{AFli} < f_{AFlj}$ and $f_{AFlj} < f_{AFhi} < f_{AFhj}$ is shown in Figure 5(c).

4. Simulation Results for Multistandard Receiver

Taking four input signals as examples, the parameters are shown in Table 1. The sampling frequency is 24 MHz. Time delay between two sampling channels is 1200 ps.

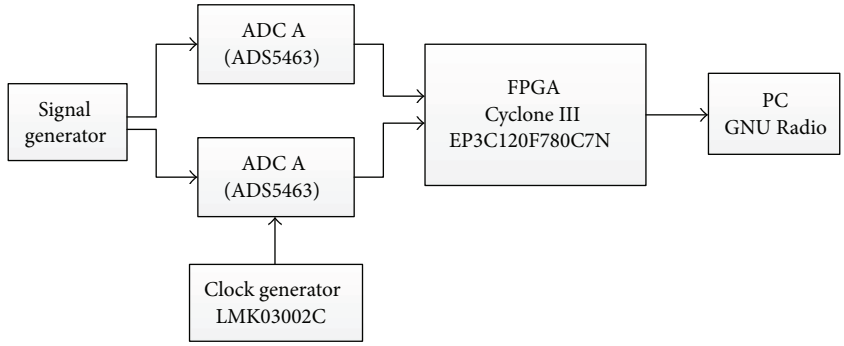


Figure 9: Structure of test platform.

After BPS, signals S_1 and S_2 overlap at around 4 MHz, and signals S_3 and S_4 overlap at around 8 MHz, as shown in Figure 6.

After applying the antialiasing filters designed in Section 2, the overlapping signals can be separated. In Figure 7, S_2 and S_4 are suppressed, and only S_1 and S_3 are left. In Figure 8, S_1 and S_3 are suppressed, and S_2 and S_4 are left. The suppressed effects are more than 30 dB.

Using further low-pass or high-pass filter design, the four signals can be separated.

Table 2: Parameters of three input signals.

Number	Central frequency (GHz)	Signal types	Zone index	Frequency offset (MHz)	Down sampled by 2.5 (MHz)
S_{1r}	2.4275	Bluetooth	101	3.5	3.5
S_{2r}	2.450	Wi-Fi	102	2	2
S_{3r}	2.4106	ZigBee	100	10.6	1

5. Hardware Test

5.1. Platform Design. In order to test in a real environment, hardware platform is designed using the structure as shown in Figure 9.

Signal generator is used to generate input RF signals. Two ADS5463 ADCs are used to realize second-order BPS. LMK03002C clock generator contributes time delay to ADC B. Antifilters are implemented in FPGA. Digital down conversion also should be implemented in FPGA. Low speed digital signals are transmitted to PC and received by GNU Radio.

5.2. Time Delay Compensation. According to (2), the theoretical phase difference between two ADC outputs can be described as

$$\theta_t = 2\pi n T_\Delta f_s \tag{16}$$

Considering the time delay error caused by hardware, practical phase difference can be written as

$$\theta_r = 2\pi n\left(T_\Delta + T^{'}\right)f_s + 2\pi\left(T_\Delta + T^{'}\right)f \tag{17}$$

where $T^{'}$ is time error which can be estimated by fitting the measured data. f is the frequency offset. $2\pi(T_\Delta + T^{'})f$ denotes the compensation for group delay. In real environment, (17) is used to design antialiasing filters.

5.3. Test Results. Taking three input signals as examples, the parameters are shown in Table 2. The sampling frequency is 24 MHz. Theoretical time delay between two sampling channels is 1200 ps, that is, clock generate is set with delay 1200 ps. By fitting the measured data, time delay error can be estimated as $T' \approx 129ps$. Then, time delay 1329 ps is used to design antialiasing filters in FPGA.

After BPS, three signals lies in 3.5 MHz, 2 MHz, and 10.6 MHz, respectively. To decrease the signal speed, signals are down sampled by 2.5, that is, signals send to PC with a sampling rate of 9.6 MHz. After down sampled, three signal lies in 3.5 MHz, 2 MHz, and 1 MHz, respectively. S_{2r} and S_{3r} are too closed to separate, so we design antialiasing filters to separate them. After applying the antialiasing filters designed in FPGA, S_{2r} and S_{3r} can be separated in channel one and channel two. In Figure 10, S_{2r} is suppressed, and only S_{3r} and S_{1r} are left. In Figure 11, S_{3r} is suppressed, and S_{2r} and S_{1r} are left.

6. Conclusions

Antialiasing filters were designed to separate more than two aliasing signals. The test results show that the antialiasing filters can suppress more than 30 dB. Based on this method, constraint conditions for selecting sampling rates are given. The algorithm is proven to be correct by hardware analysis. Compared with existing receivers in telemedicine monitoring systems, the proposed structure can solve the problem of aliasing and realize a multistandard receiver that can receive different standard signals simultaneously. This receiver can significantly improve spectrum utilization as well as the flexibility of receivers for telemedicine monitoring systems.

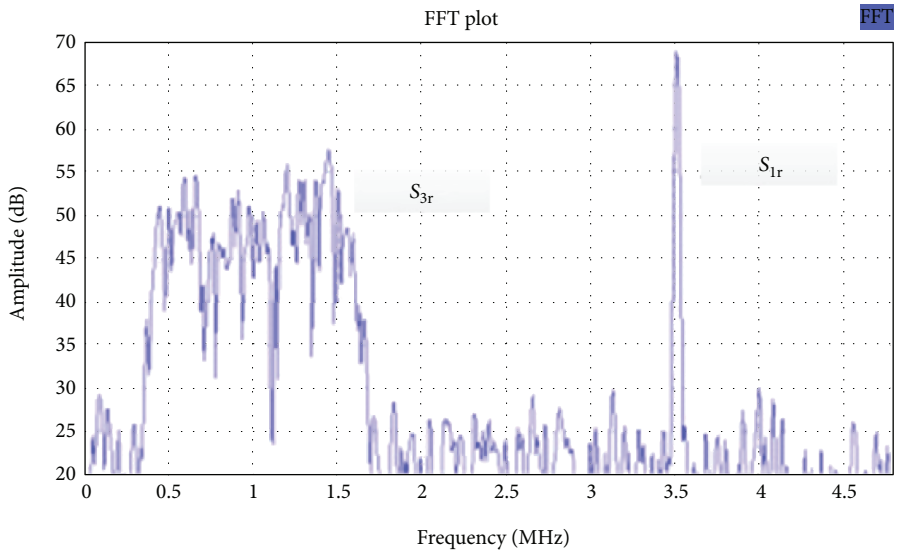


Figure 10: Signal spectra received by GNU Radio (channel one).

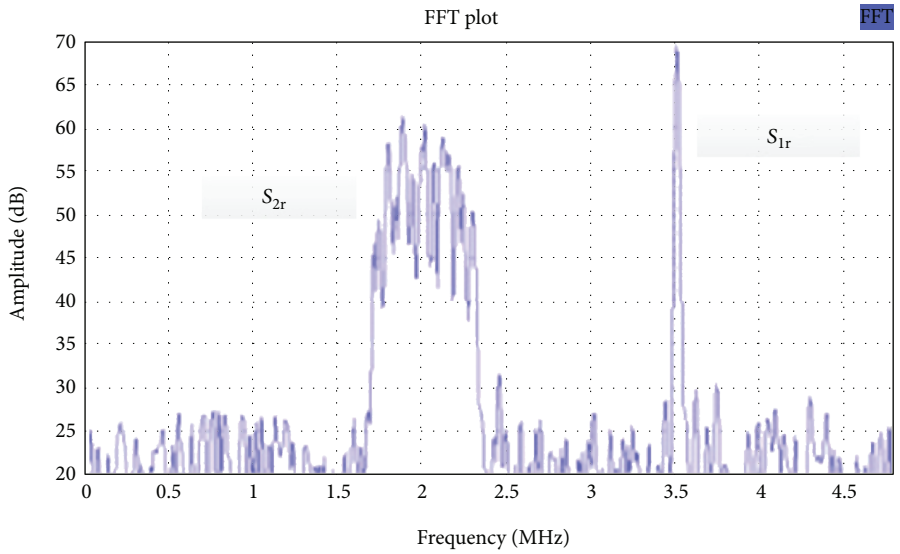


Figure 11: Signal spectra received by GNU Radio (channel two).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (61601464 and 61771474) and in part by the Fundamental Research Funds for the Central Universities (2013QNA49).

References

[1] M. Aguas, J. Del Hoyo, R. Faubel, B. Valdivieso, and P. Nos, "Telemedicine in the treatment of patients with inflammatory bowel disease," *Gastroenterología y Hepatología*, vol. 40, no. 9, pp. 641–647, 2017.

[2] N. Jiang, Y. Zhuang, and D. K. W. Chiu, "Multiple transmission optimization of medical images in recourse-constraint mobile telemedicine systems," *Computer Methods and Programs in Biomedicine*, vol. 145, pp. 103–113, 2017.

[3] M. Cao, J. Li, Y. Dai, F. Yin, and K. Xu, "Multiband phase-modulated RoF link with coherent detection and bandpass sampling," *IEEE Photonics Technology Letters*, vol. 27, no. 21, pp. 2308–2311, 2015.

[4] J. Baek, "Design of second-order BPS systems for the cancellation of multiple aliasing," *Journal of the Institute of Electronics and Information Engineers*, vol. 52, no. 3, pp. 162–170, 2015.

[5] M. A. Messaoud, F. M. Ghannouchi, R. Barrak, and A. Ghazel, "SDR based multi-band subsampling receivers for GNSS applications," in *2014 1st International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, pp. 497–501, Sousse, Tunisia, 2014.

[6] J.-M. Munoz-Ferreras, R. Gomez-Garcia, and F. Perez-Martinez, "RF front-end concept and implementation for direct sampling of multiband signals," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 58, no. 3, pp. 129–133, 2011.

[7] M. Choe and K. Kim, "Bandpass sampling algorithm with normal and inverse placements for multiple RF signals," *IEICE Transactions on Communications*, vol. E88-B, no. 2, pp. 754–757, 2005.

[8] S. Sarwana, D. E. Kirichenko, V. V. Dotsenko, A. F. Kirichenko, S. B. Kaplan, and D. Gupta, "Multi-band digital-RF receiver," *IEEE Transactions on Applied Superconductivity*, vol. 21, no. 3, pp. 677–680, 2011.

[9] H.-m. Wang, J.-h. Kim, F.-g. Wang, S.-h. Lee, and X.-s. Wang, "Design of BPS digital frontend for software defined radio receiver," *Journal of Central South University*, vol. 22, no. 12, pp. 4709–4716, 2015.

Hindawi
Journal of Sensors
Volume 2018, Article ID 5738274, 9 pages
https://doi.org/10.1155/2018/5738274

Research Article

T-S Fuzzy-Based Optimal Control for Minimally Invasive Robotic Surgery with Input Saturation

Faguang Wang[1,2] **Hongmei Wang**[1] **Yong Zhang**[1] **and Xijin Guo**[1]

[1]*School of Information and Control Engineering, China University of Mining and Technology, No. 1 Daxue Road, Xuzhou 221116, China*
[2]*Collaborative Innovation Center of Intelligent Mining Equipment, China University of Mining and Technology, No. 1 Daxue Road, Xuzhou 221116, China*

Correspondence should be addressed to Hongmei Wang; whm99@cumt.edu.cn

Received 23 December 2017; Accepted 25 March 2018; Published 23 April 2018

Academic Editor: Mucheol Kim

A minimally invasive surgery robot is difficult to control when actuator saturation exists. In this paper, a Takagi-Sugeno fuzzy model-based controller is designed for a minimally invasive surgery robot with actuator saturation, which is difficult to control. The contractively invariant ellipsoid theorem is applied for the actuator saturation. The proposed scheme can be derived using the H-infinity control theorem and parallel distributed compensation. The result is rebuilt in the form of linear matrix inequalities for easier calculation by computer. Meanwhile, the uniformly ultimately bounded stable and the prescribed H-infinity control performance can be guaranteed. The proposed scheme is simulated in a Novint Falcon haptic device system.

1. Introduction

In minimally invasive robotic surgery, the work space is limited precisely. The mechanical structure and the electrical characteristics can be additional constraint boundaries for system inputs and outputs. Therefore, the controller of the surgery robot has a rigid input saturation requirement. However, if input saturation occurs, the output performance of the controlled object cannot satisfy the designed requirement, which can result in the decline of the closed-loop system response. The output overshoot cannot be suppressed well, even becoming unstable [1, 2]. This situation is prohibited in a minimally invasive robotic surgery.

Many researchers investigated the input saturation problem and provided some solutions. Buckley [3] proposed an anti-reset windup method for the integral saturation problem. The error between the controller output and object input was used as a complementary feedback for the controlled system. Hanus et al. [4] proposed the condition technique, where the controller output continuously tracks for a new reference input and is located out of the saturation area to avoid the saturation case. The pole-placement method was applied to the antiwindup control, which assigns the poles of nonlinear system in the desired disk for stable analysis in [5]. Although the variable structure antiwindup controller showed a good performance in the integrator windup case [6], some preset parameters were needed that were difficult to determine because expert knowledge was required. Some antiwindup controllers were also based on observer [7], internal model control [8], saturation feedback control [9], and dynamic complement [10].

The T-S fuzzy theorem, which is an important part of the fuzzy control theorem, was proposed by Takagi and Sugeno in 1985 [11]. It is utilized in both system stable control and

model identification [11, 12]. Specifically, it has a good performance in nonlinear system control [12]. The T-S fuzzy model, which is the summation of the product of the T-S fuzzy local models and their corresponding membership functions, can approximate the nonlinear system under an arbitrary degree of accuracy [12]. The T-S fuzzy model is nonlinear, and its controller design is difficult. Consequently, a procedure called parallel distributed compensation (PDC) makes the T-S fuzzy controller design easier. PDC was proposed by Wang et al. in 1995 [13]. The fuzzy sets of the PDC controller are similar with the fuzzy model. Under each controller rule, a controller can be designed for the local T-S model. By summation of the product of these controllers and their corresponding membership functions, the total controller of a nonlinear system can be represented. System stability is usually considered by the Lyapunov function. Although PDC provides a design procedure for the T-S fuzzy controller, the calculation of controller gains for global stability is still difficult, particularly in the presence of many controller rules. A numerical optimization method called linear matrix inequalities (LMIs) solved this kind of problem. LMI was defined by Willems in 1971 [14]. Neestorv and Nemirovskii [15] proposed an interior point method, which can directly solve the LMI convex optimization problem. LMI was first applied to the T-S fuzzy system stabilization analysis in [16]. In the succeeding years, LMI became the focus of increasing number of researchers [17]. Except for a few special cases without analytical solution, LMIs are generally efficient [11]. Furthermore, the LMI toolbox software provides a direct shortcut to the computer solution of LMIs. Because PDC and LMIs provide a better T-S fuzzy system controller design, they are applied in the proposed controller design in this paper.

For the advantage of the T-S fuzzy theorem, it was widely used to deal with the actuator saturation problem. In [18], based on the T-S fuzzy model, a robust dissipative controller was designed for the multiple-input multiple-output (MIMO) system with saturated time-delay input and parameter uncertainty. Their results show that the closed-loop system can be stable, but the stabilization time cannot be guaranteed. For this problem, [19, 20] provided a finite-time control by optimal control and estimated the attraction domain. The combination of the T-S fuzzy model and optimal control show an animated controller design of a nonlinear system with actuator saturation [21]. For optimal control, Hu et al. [22] provided a useful control method for the saturation system with actuator saturation based on a contractively invariant ellipsoid. This study was continued in her work by BMIs [8]. Consequently, many researchers focused on actuator saturation problems, where the T-S fuzzy based controller design method was popular. The Lyapunov stability criterion-based PDC fuzzy controller was designed for the actuator saturation system [23, 24]. Many researchers contributed in completing this theorem [25, 26]. In this paper, the result in Hu et al. [22] was adopted and rebuilt into a set of LMIs. Meanwhile, a predetermined H_∞ norm was satisfied.

This paper is organized as follows. The basic T-S fuzzy theorem is presented in Section 2. The proposed solution for the actuator saturation problem is presented in Section 3. The simulation is presented in Section 4. The paper concludes in Section 5.

2. General T-S Fuzzy Model and Control

In this paper, the considered nonlinear system with disturbance $W(t)$ is

$$\dot{X}(t) = f(X(t)) + g(X(t))U(t) + W(t), \tag{1}$$

where $X(t) \in R^{n\times 1}$, $f(X(t)) \in R^{n\times 1}$, $g(X(t)) \in R^{n\times m}$, $U(t) \in R^{m\times 1}$, and $\|W(t)\| \le W_b$ and W_b is the boundary of disturbance. In the T-S fuzzy model, the premise variables of the T-S fuzzy rules must be measurable, and they can represent some properties of the nonlinear system. Therefore, a suitable selection of these variables is very important for the accuracy and reliability of the T-S fuzzy model. When the T-S fuzzy premise variables have been defined, a corresponding local model can be provided by these variable values. The ith rule of the T-S fuzzy system for a nonlinear system is as follows:

The ith rule: If $z_1(t)$ is M_{i1}, …, and $z_p(t)$ is M_{ip}, then

$$\dot{X}(t) = A_i X(t) + B_i U(t), \tag{2}$$

where $i = 1, 2, \dots, L$ is the rule number, $\mathbf{A}_i \in R^{n\times n}$ and $\mathbf{B}_i \in R^{n\times m}$ are the local model parameter matrices, $z_1, z_2, \dots, z_p$ are the premise variables, and $M_{i1}, \dots, M_{ip}$ are the fuzzy sets. The nonlinear system can be approximated by the overall T-S fuzzy system

$$\dot{X}(t) = \sum_{i=1}^{L} h_i(Z(t))\{A_i X(t) + B_i U(t)\}, \tag{3}$$

where $Z(t) = [z_1(t) z_2(t) \cdots z_p(t)]$ and $h_i(Z(t)) = \mu_i(Z(t))/\sum_{i=1}^{L}\mu_i(Z(t))$, $\mu_i(Z(t)) = \prod_{j=1}^{p}\theta_{ij}(z_j(t))$ in which $\theta_{ij}(Z(t))$ is the grade of the membership of $z_j(t)$ in M_{ij}. Note that the membership function should satisfy the following equation:

$$\begin{aligned} h_i(Z(t)) &\ge 0, \\ \sum_{i=1}^{L} h_i(Z(t)) &= 1. \end{aligned} \tag{4}$$

Currently, the nonlinear system of (1) is changed to the T-S fuzzy model (3). Because the local T-S model of (2) is linear, its feedback controller can be designed easily as

$$U_j = K_j X(t), \tag{5}$$

where $K_j \in R^{m\times n}$ and $j = 1, 2, \dots, L$. By PDC and similar fuzzy rules of (2), the overall T-S fuzzy controller is designed as

$$U(t) = \sum_{j=1}^{L} h_j(Z(t)) K_j X(t). \tag{6}$$

3. H_∞ Robust Controller Design for Input Saturation

The traditional T-S fuzzy theorem cannot provide results in the control problem for a nonlinear system with input saturation. In this study, the proposed method can solve the input saturation problem. The controlled closed-loop system can be uniformly ultimately bounded (UUB) stable, and the prescribed H-infinity norm can be guaranteed. Furthermore, the result is shown in LMIs, which can be directly solved by programming.

Considering the saturation and uncertainty, the nonlinear system of (1) can be

$$\dot{X}(t) = f(X(t)) + g(X(t)) \cdot \mathrm{Sat}(U(t)) + W(t), \tag{7}$$

where

$$\mathrm{Sat}(U(t)) = \begin{cases} \mathrm{sgn}(U(t)) \cdot b, & \|U(t)\| > b, \\ U(t), & \|U(t)\| \le b, \end{cases} \tag{8}$$

$b > 0$, is the boundary value of the saturation function, and the other symbols are similar with (1). Following (3), the T-S fuzzy model can be

$$\dot{X}(t) = \sum_{i=1}^{L} h_i [A_i X(t) + B_i \mathrm{Sat}(U(t))], \tag{9}$$

where

$$U(t) = \sum_{j=1}^{L} h_j(Z(t)) U_j(t). \tag{10}$$

Based on the work of Hu et al. [1], for the feedback controller, $U_j(t) = K_j X(t)$, if the ellipsoid $\varepsilon(P, 1)$ is contractively invariant, and matrixes $\mathbf{P} > 0$, $\mathbf{K}_i$, and $\mathbf{H}$ ($m \times n$ matrix) make the initial value of $X_0 \in \varepsilon(P, 1)$ and

$$\varepsilon(P, 1) \subset \ell(H), \tag{11}$$

where the polyhedron $\ell(H) \coloneqq \{X(t) \in R^n : |H_j X(t)| \le 1, j \in [1, m]\}$ and H_j is the jth row of H. The saturation feedback controller $\mathrm{Sat}(U(t))$ satisfies

$$\begin{aligned} \mathrm{Sat}\left(U_j(t)\right) = \mathrm{Sat}\left(K_j X(t)\right) \in \mathrm{co} \\ \cdot \left\{D_r K_j X(t) + D_r^- H_j X(t), r \in [1, 2^m]\right\}, \end{aligned} \tag{12}$$

where D_r is an m-by-m diagonal matrix (diagonal elements are 0 or 1) and $D_r^- = I - D_r$ with $r \in [1, 2^m]$.

If it is only applied to the T-S fuzzy local model, the global stability of the nonlinear system cannot be satisfied. For global stability, some changes are needed. By (7) and (9), the nonlinear system with disturbance is as follows:

$$\begin{aligned} \dot{X}(t) &= f(X(t)) + g(X(t))\mathrm{Sat}(U(t)) + W(t) \\ &= \sum_{i=1}^{L} h_i(Z(t)) \cdot (A_i X(t) + B_i \mathrm{Sat}(U(t))) + f(X(t)) \\ &\quad - \sum_{i=1}^{L} h_i(Z(t)) A_i \cdot X(t) \\ &\quad + \left(g(X(t)) - \sum_{i=1}^{L} h_i(Z(t)) B_i\right) \mathrm{Sat}(U(t)) + W(t). \end{aligned} \tag{13}$$

Equation (12) can be guaranteed if each D_r satisfies the following inequality:

$$\begin{aligned} \dot{X} &\le \sum_{i=1}^{L} h_i(Z(t)) \left(A_i X(t) + B_i \sum_{j=1}^{L} h_j(Z(t)) \left(D_r K_j X(t) + D_r^- H_j X(t)\right)\right) \\ &\quad + \left(f(X(t)) - \sum_{i=1}^{L} h_i(Z(t)) A_i X(t)\right) \\ &\quad + \left(g(X(t)) - \sum_{i=1}^{L} h_i(Z(t)) B_i\right) \\ &\quad \cdot \left(\sum_{j=1}^{L} h_j(Z(t)) \left(D_r K_j X(t) + D_r^- H_j X(t)\right)\right) + W(t) \\ &= \sum_{i=1}^{L} \sum_{j=1}^{L} h_i(Z(t)) h_j(Z(t)) \left(A_i + B_i \left(D_r K_j + D_r^- H_j\right)\right) X(t) \\ &\quad + \Delta f + \Delta g + W(t), \end{aligned} \tag{14}$$

where

$$\begin{aligned} \Delta f &= \left(f(X(t)) - \sum_{i=1}^{L} h_i(Z(t)) A_i X(t)\right), \\ \Delta g &= \sum_{i=1}^{L} h_i(Z(t)) \sum_{j=1}^{L} h_j(Z(t)) (g(X(t)) - B_i) \\ &\quad \cdot \left(D_r K_j + D_r^- H_j\right) X(t). \end{aligned} \tag{15}$$

Some $\|\delta_{Ai}\| \le 1$, $\|\delta_{Bi}\| \le 1$, A_t, and B_t can be found and satisfy the following inequality:

$$\begin{aligned} \|\Delta f\| &= \left\| \sum_{i=1}^{L} h_i(Z(t)) \Delta A_i X(t) \right\| \le \left\| \sum_{i=1}^{L} h_i(Z(t)) \delta_{Ai} A_t X(t) \right\| \\ &\le \|A_t X(t)\|, \\ \|\Delta g\| &\le \left\| \sum_{j=1}^{L} h_j(Z(t)) B_t \left(D_r K_j + D_r^- H_j\right) X(t) \right\|, \end{aligned} \tag{16}$$

where $\Delta A_i = \delta_{Ai} A_t$ and $\Delta B_i = \delta_{Bi} B_t$.

H_∞ control performance is defined as

$$\frac{\int_0^{t_f} X^T(t) Q X(t) dt}{\int_0^{t_f} W^T(t) W(t) dt} < \rho, \tag{17}$$

where t_f is the terminal time, $\mathbf{Q}$ is a positive-definite matrix, and ρ is the prescribed H_∞ norm, which is greater than 0 and less than 1. If ρ is minimized, the effect of $W(t)$ on $X(t)$ is minimized. Considering the initial condition, (17) can be changed as

$$\int_0^{t_f} X^T(t)QX(t)dt < X^T(0)PX(0) + \rho\int_0^{t_f} W^T(t)W(t)dt, \quad (18)$$

where $\mathbf{P}$ is some symmetric positive-definite weighting matrix. Set the Lyapunov function as

$$V(t) = X^T(t)PX(t), \quad (19)$$

where $\mathbf{P} > 0$. For simplicity, (t) is omitted in the following sections. Then, the following theorem can be obtained.

Theorem 1. *Under condition (11), if controller (12) is applied to a nonlinear system (7) and there exists a positive-definite matrix* $\mathbf{P} > 0$*, such that the following matrix inequalities*

$$\begin{aligned} &PA_i + A_i^T P + PB_i\left(D_rK_j + D_r^-H_j\right) + \left(D_rK_j + D_r^-H_j\right)^T \\ &\quad \cdot B_i{}^T P + A_t{}^T A_t + \left(D_rK_j + D_r^-H_j\right)^T B_t{}^T B_t\left(D_rK_j + D_r^-H_j\right) \\ &\quad + 2PP + \rho^{-2}PP + Q_i < 0 \end{aligned} \quad (20)$$

are satisfied for each i, j, and r, then the closed-loop system is UUB, and the H_∞ *control performance (18) is guaranteed as prescribed* ρ^2*.*

Proof 1. Using (14), the derivative of $V(t)$ is

$$\begin{aligned} \dot V(t) &= X^T P\dot X + \dot X^T PX = X^T P\left\{\sum_{I=1}^{l} h_i[A_iX + B_i\mathrm{Sat}(U) + W]\right\} \\ &\quad + \left\{\sum_{i=1}^{l} h_i[A_iX + B_i\mathrm{Sat}(U) + W]\right\}^T PX \le X^TP \\ &\quad \cdot \left\{\sum_{i=1}^{L}\sum_{j=1}^{L} h_ih_j\left(A_i + B_i\left(D_rK_j + D_r^-H_j\right)\right)X + \Delta f + \Delta g + W\right\} \\ &\quad + \left\{\sum_{i=1}^{L}\sum_{j=1}^{L} h_ih_j\left(A_i + B_i\left(D_rK_j + D_r^-H_j\right)\right)X + \Delta f + \Delta g + W\right\}^T \\ &\quad \cdot PX \le \sum_{i=1}^{L}\sum_{j=1}^{L} h_ih_j\Big\{X^T\Big(P\left(A_i + B_i\left(D_rK_j + D_r^-H_j\right)\right) \\ &\qquad + \left(A_i + B_i\left(D_rK_j + D_r^-H_j\right)\right)^T P\Big)X \\ &\qquad + (\Delta f)^T(\Delta f) + X^TPPX + (\Delta g)^T(\Delta g) \\ &\qquad + X^TPPX\Big\} + X^TPW + W^TPX. \end{aligned} \quad (21)$$

Using (16), (21) can be

$$\begin{aligned} \dot V &\le \sum_{i=1}^{L}\sum_{j=1}^{L} h_ih_j\Big\{X^T\Big(P\left(A_i + B_i\left(D_rK_j + D_r^-H_j\right)\right) \\ &\qquad + \left(A_i + B_i\left(D_rK_j + D_r^-H_j\right)\right)^T P\Big)X + (A_tX)^T(A_tX) \\ &\qquad + X^TPPX + \left(B_t\left(D_rK_j + D_r^-H_j\right)X\right)^T \\ &\qquad \cdot\left(B_t\left(D_rK_j + D_r^-H_j\right)X\right) + X^TPPX\Big\} + X^TPW \\ &\qquad + W^TPX \\ &= \sum_{i=1}^{L}\sum_{j=1}^{L} h_ih_jX^T\Big\{\Big(PA_i + A_i^TP + PB_i\left(D_rK_j + D_r^-H_j\right) \\ &\qquad + \left(D_rK_j + D_r^-H_j\right)^T B_i{}^TP + A_t{}^TA_t + \left(D_rK_j + D_r^-H_j\right)^T \\ &\qquad \cdot B_t{}^TB_t\left(D_rK_j + D_r^-H_j\right) + 2PP\Big)\Big\}X - \left(\rho^{-1}PX - \rho W\right)^T \\ &\qquad \cdot\left(\rho^{-1}PX - \rho W\right) + \rho^2W^TW + \rho^{-2}X^TPPX \\ &\le \sum_{i=1}^{L}\sum_{j=1}^{L} h_ih_jX^T\Big\{\Big(PA_i + A_i^TP + PB_i\left(D_rK_j + D_r^-H_j\right) \\ &\qquad + \left(D_rK_j + D_r^-H_j\right)^T B_i{}^TP + A_t{}^TA_t \\ &\qquad + \left(D_rK_j + D_r^-H_j\right)^T B_t{}^TB_t\cdot\left(D_rK_j + D_r^-H_j\right) \\ &\qquad + +2PP\Big)\Big\}X + \rho^2W^TW + \rho^{-2}X^TPPX. \end{aligned} \quad (22)$$

Under the ith rule, if there is

$$\frac{\int_0^{t_f} X^T(t)\left(h_i^2Q_i\right)X(t)dt}{\int_0^{t_f} W^T(t)W(t)dt} \le \frac{\int_0^{t_f} X^T(t)Q_iX(t)dt}{\int_0^{t_f} W^T(t)W(t)dt} < \rho, \quad (23)$$

where $Q_i > 0$ and $\sum_{i=1}^{L}(h_iQ_i) > 0$, there is

$$\frac{\int_0^{t_f} X^T(t)\left(\sum_{i=1}^{L} h_iQ_i\right)X(t)dt}{\int_0^{t_f} W^T(t)W(t)dt} < \rho \quad (24)$$

for the nonlinear system. If $\sum_{i=1}^{L}(h_iQ_i) = Q$, (24) can be the H_∞ control performance defined in (17). Therefore, for each simulation time, the H_∞ control performance defined as (17) can be guaranteed. Then, in the total simulation, a prescribed H_∞ norm can be guaranteed. Meanwhile, (18) can be as follows:

$$\int_{t=0}^{t=t_f}\left(X^T\left(\sum_{i=1}^{L} h_iQ_i\right)X\right)dt < X^T(0)PX(0) + \rho\int_{t=0}^{t=t_f}\left(W^TW\right)dt. \quad (25)$$

If

$$\begin{aligned} &PA_i + A_i^TP + PB_i\left(D_rK_j + D_r^-H\right) + \left(D_rK_j + D_r^-H\right)^T B_i{}^TP \\ &\quad + A_t{}^TA_t + \left(D_rK_j + D_r^-H\right)^T B_t{}^TB_t\left(D_rK_j + D_r^-H\right) \\ &\quad + 2PP < -\rho^{-2}PP - Q_i, \end{aligned} \quad (26)$$

there is

$$\dot{V} < \sum_{i=1}^{L}\sum_{j=1}^{L} h_i h_j X^T \left\{-\rho^{-2}PP - Q_i\right\} X + \rho^2 W^T W + \rho^{-2} X^T PPX$$
$$= X^T \left\{-\rho^{-2}PP - \left(\sum_{i=1}^{L} h_i Q_i\right)\right\} X + \rho^2 W^T W$$
$$+ \rho^{-2} X^T PPX = -X^T \left(\sum_{i=1}^{L} h_i Q_i\right) X + \rho^2 W^T W. \quad (27)$$

By integrating (27), the following result can be obtained:

$$V(t_f) - V(0) < \int_{t=0}^{t=t_f} \left(-X^T \left(\sum_{i=1}^{L} h_i Q_i\right) X + \rho^2 W^T W\right) dt. \quad (28)$$

The Lyapunov function (19) can be 0 when the system is stable. Therefore,

$$\int_{t=0}^{t=t_f} \left(X^T \left(\sum_{i=1}^{L} h_i Q_i\right) X\right) dt < X^T(0) PX(0) + \rho^2 \int_{t=0}^{t=t_f} \left(W^T W\right) dt. \quad (29)$$

The above equation is the H_∞ performance defined in (25), which satisfies (18). Note that the H_∞ norm is changed to ρ^2.

End of proof.

In Theorem 1, it is difficult to find the solution for (20). Therefore, the transformation of inequality (20) into LMIs is necessary. By setting $R = P^{-1}$ and left and right multiplication of R, (20) can be

$$A_i R + R A_i^T + B_i \left(D_r K_j + D_r^- H_j\right) R + R \left(D_r K_j + D_r^- H_j\right)^T B_i^T$$
$$+ R {A_t}^T A_t R + R \left(D_r K_j + D_r^- H_j\right)^T {B_t}^T B_t \left(D_r K_j + D_r^- H_j\right) R$$
$$+ 2 + \rho^{-2} + R Q_i R < 0. \quad (30)$$

Let $\hat{K}_i = K_i \cdot R$, $\hat{H}_i = H_i \cdot R$, and $\hat{Q}_i = \left({A_t}^T A_t + Q_i\right)^{-1}$. By Schur complements, (30) can be changed as

$$\begin{bmatrix} A^* & \left(B_t \left(D_r \hat{K}_j + D_r^- \hat{H}_j\right)\right)^T & R \\ B_t \left(D_r \hat{K}_j + D_r^- \hat{H}_j\right) & -I & 0 \\ R & 0 & -\hat{Q}_i \end{bmatrix} < 0, \quad (31)$$

where $A^* = A_i R + R A_i^T + B_i (D_r \hat{K}_j + D_r^- \hat{H}_j) + (D_r \hat{K}_j + D_r^- \cdot \hat{H}_j)^T B_i^T + (2 + \rho^{-2}) I$, R, $\hat{K}_j$, and $\hat{H}_j$ are unknown. Condition (11) is equal to

$$\begin{bmatrix} 1 & \hat{H}_{ik} \\ {\hat{H}_{ik}}^T & R \end{bmatrix} \geq 0, \quad (32)$$

Figure 1: Novint Falcon.

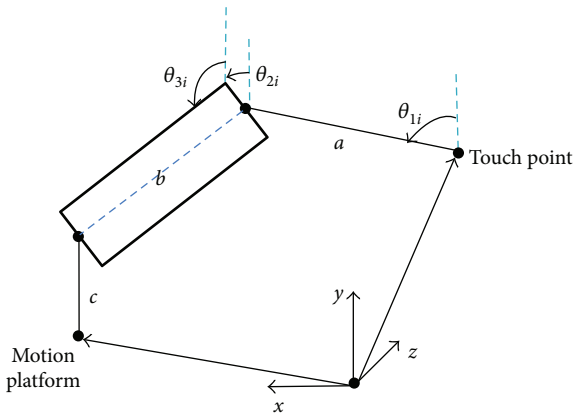


Figure 2: Mathematical model of the single chain.

where $\hat{H}_{ik}$ is the kth row of $\hat{H}_i$. Therefore, the following theorem is given.

Theorem 2. *For nonlinear system (7) with controller (12), if there exists*

$$\min \quad \rho^2,$$
$$\text{subject to} \quad \begin{bmatrix} A^* & \left(B_t \left(D_r \hat{K}_j + D_r^- \hat{H}_j\right)\right)^T & R \\ B_t \left(D_r \hat{K}_j + D_r^- \hat{H}_j\right) & -I & 0 \\ R & 0 & -\hat{Q}_i \end{bmatrix} < 0,$$
$$\begin{bmatrix} 1 & \hat{H}_{ik} \\ {\hat{H}_{ik}}^T & R \end{bmatrix} \geq 0,$$
$$R > 0, \quad (33)$$

then the closed-loop system is UUB and the H_∞ control performance (18) is guaranteed as prescribed ρ^2.

By this theorem, the H_∞ optimization problem can be changed to a constrained optimization problem by LMIs.

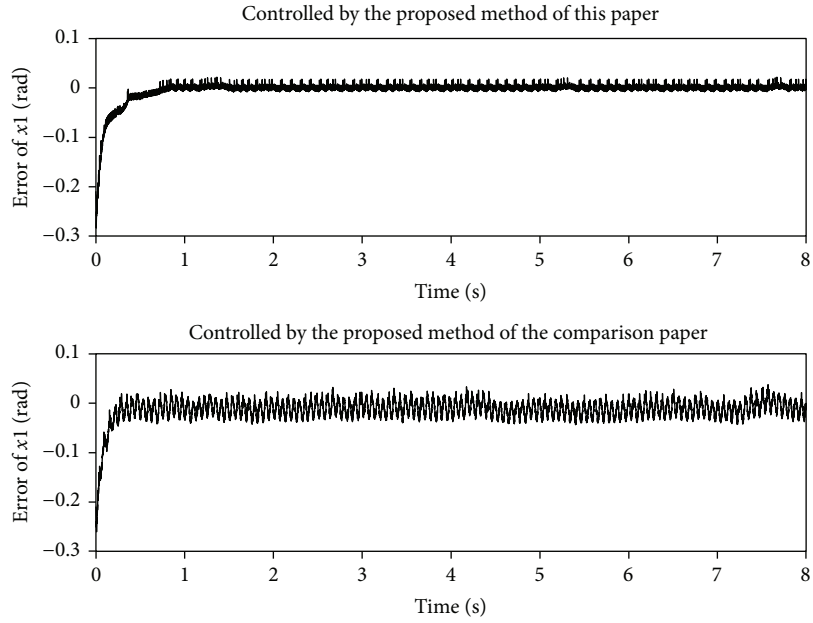


FIGURE 3: The error of x_1.

4. Numerical Example

In minimally invasive surgery, haptic tracking is very important. Novint Falcon is a simple arm robot with force feedback ability, which benefited from parallel architecture, and can be applied in minimally invasive surgery. The Novint Falcon structure is shown in Figure 1.

The Novint Falcon consists of the base, motion platform, and 3 chains. The single chain can be abstract presented in Figure 2.

Based on [27] and ignoring some nonessentials, the ith chain dynamic equation is

$$\tau = D(\theta)\ddot{\theta} + H\left(\theta, \dot{\theta}\right)\dot{\theta} + G(\theta), \tag{34}$$

where $\tau = [\tau_1, \tau_2, \tau_3]^T$ is the moment of driving force for each link; $\theta = [\theta_{1i}, \theta_{2i}, \theta_{3i}]^T$ is shown in Figure 2; $\mathbf{D}(\theta) = I_T + (J^T)^{-1}(3m_b + m_c)J^{-1}$, where I_T is the moment of inertia and m_b and m_c are masses of links b and c; $H(\theta, \dot{\theta}) = D_{\mathrm{f}} + (J^T)^{-1}(3m_b + m_c)((J^T)^{-1})'$, where D_{f} denotes the damped coefficient; $G(\theta) = -M_{\mathrm{g}}$ is the gravity moment;

$$\mathbf{J} = \begin{bmatrix} J_{11} & J_{12} & J_{13} \\ J_{21} & J_{22} & J_{23} \\ J_{31} & J_{32} & J_{33} \end{bmatrix}^{-1} \begin{bmatrix} \alpha_1 & 0 & 0 \\ 0 & \alpha_2 & 0 \\ 0 & 0 & \alpha_3 \end{bmatrix} \tag{35}$$

is the Jacobian matrix; $J_{i1} = \cos(\theta_{2i})\sin(\theta_{3i})\cos(\phi_i) - \cos(\theta_{3i})\sin(\phi_i)$, $J_{i2} = \cos(\theta_{3i})\cos(\phi_i) + \cos(\theta_{2i})\sin(\theta_{3i})\sin(\phi_i)$, $J_{i3} = \sin(\theta_{2i})\sin(\theta_{3i})$, $\alpha_i = \sin(\theta_{2i} - \theta_{1i})\sin(\theta_{3i})$, and ϕ are the angle between the chain and the coordinate frame (x, y, z). Because the exact parameters are difficult to determine, set $\mathbf{I}_T$ and $\mathbf{M}_{\mathrm{g}}$ as identity matrices and $3m_b + m_c = 1$ in this example. This is reasonable in this paper because the error between the actual value and the required value can be distributed in the disturbance or parameter uncertainty for this system. Therefore, the matrix $\mathbf{D}(\theta)$ is invertible. Set $\theta_{1i} = x_1(t)$, $\theta_{2i} = x_2(t)$, $\theta_{3i} = x_3(t)$, $\tau = u(t)$, and

$$\begin{aligned} \dot{x}_1(t) &= x_4(t), \\ \dot{x}_2(t) &= x_5(t), \\ \dot{x}_3(t) &= x_6(t). \end{aligned} \tag{36}$$

The θ of the other chains are set to be closed to require values, and the errors are distributed in the disturbance. Therefore, the system (34) can be as follows:

$$\begin{bmatrix} \dot{x}_4(t) \\ \dot{x}_5(t) \\ \dot{x}_6(t) \end{bmatrix} = A \begin{bmatrix} x_4(t) \\ x_5(t) \\ x_6(t) \end{bmatrix} + Bu(t) + E, \tag{37}$$

where $A = -D(\theta)^{-1}H(\theta, \dot{\theta})$, $B = D(\theta)^{-1}$, and $E = -D(\theta)^{-1}G(\theta)$. Considering actuator saturation, the corresponding ith fuzzy rule is as follows:

Rule i. If A is A_i and B is B_i, then (37) is

$$\begin{bmatrix} \dot{x}_4(t) \\ \dot{x}_5(t) \\ \dot{x}_6(t) \end{bmatrix} = A_i \begin{bmatrix} x_4(t) \\ x_5(t) \\ x_6(t) \end{bmatrix} + B_i \mathrm{Sat}(u(t)), \tag{38}$$

where $A_1 = A(\max(\theta_{1i}), \max(\theta_{2i}), \max(\theta_{3i}))$, $A_2 = A(\min(\theta_{1i}), \max(\theta_{2i}), \max(\theta_{3i}))$, $A_3 = A(\max(\theta_{1i}), \min(\theta_{2i}), \max(\theta_{3i}))$, $A_4 = A(\max(\theta_{1i}), \max(\theta_{2i}), \min(\theta_{3i}))$, $A_5 = A(\min(\theta_{1i}), \min(\theta_{2i}), \max(\theta_{3i}))$, $A_6 = A(\min(\theta_{1i}), \max(\theta_{2i}), \min(\theta_{3i}))$, $A_7 =$

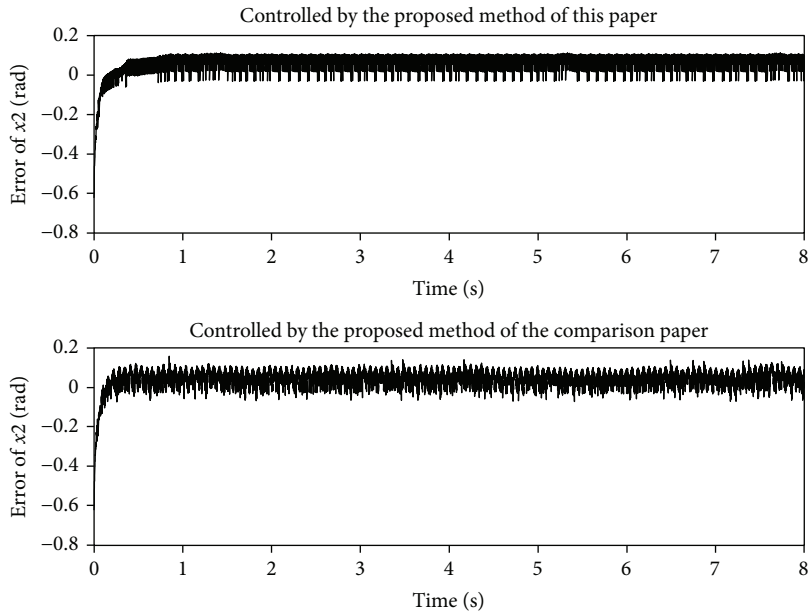


FIGURE 4: The error of x_2.

$A(\max(\theta_{1i}), \min(\theta_{2i}), \min(\theta_{3i}))$, and $A_8 = A(\min(\theta_{1i}), \min(\theta_{2i}), \min(\theta_{3i}))$; B_i is calculated similarly, and E is treated as a disturbance. Although A_i depends on θ, the values of $\max(\theta)$ and $\min(\theta)$ do not depend on the system input. Therefore, they can be used for the fuzzy rule definition. Set $\theta_{1i} \in [0, 0.7854]$, $\theta_{2i} \in [0, 0.7854]]$, and $\theta_{3i} \in [1.9199, 2.2689]$. $\mathrm{Sat}(u(t)) = \mathrm{sgn}(u(t)) \cdot \min\{|u(t)|, 10\}$. The H_∞ norm is 0.3. The reference value is

$$[x_{1r} \;\; x_{2r} \;\; x_{3r} \;\; x_{4r} \;\; x_{5r} \;\; x_{6r}] = [0.7 \quad 0.7 \quad 2.2 \quad 0 \quad 0 \quad 0]. \tag{39}$$

Using Theorem 2, the error system feedback controller gains can be calculated as follows:

$$\begin{aligned}
\mathrm{K1} &= \left[-2.9 \times 10^4, -1.8 \times 10^5, 2.6 \times 10^4 ; 3.0 \times 10^4, 6.5 \times 10^4, -3.4 \times 10^4 ; -7.9 \times 10^2, 1.2 \times 10^5, 7.7 \times 10^3\right],\\
\mathrm{K2} &= \left[1.8 \times 10^4, 1.6 \times 10^5, -1.4 \times 10^4 ; -1.8 \times 10^4, -3.2 \times 10^4, 2.1 \times 10^4 ; -1.2 \times 10^2, -1.3 \times 10^5, -6.9 \times 10^3\right],\\
\mathrm{K3} &= \left[-3.8 \times 10^4, -4.7 \times 10^5, 2.2 \times 10^4 ; 1.4 \times 10^4, 1.1 \times 10^5, -1.2 \times 10^4 ; 2.4 \times 10^4, 3.6 \times 10^5, -1.1 \times 10^4\right],\\
\mathrm{K4} &= \left[1.9 \times 10^4, 2.9 \times 10^5, -7.6 \times 10^3 ; -1.7 \times 10^4, 2.2 \times 10^3, 2.2 \times 10^4 ; -1.4 \times 10^3, -2.9 \times 10^5, -1.4 \times 10^4\right],\\
\mathrm{K5} &= \left[-2.0 \times 10^4, -2.4 \times 10^5, 1.2 \times 10^4 ; 9.7 \times 10^3, 5.1 \times 10^4, -9.5 \times 10^3 ; 9.8 \times 10^3, 1.8 \times 10^5, -2.1 \times 10^3\right],\\
\mathrm{K6} &= \left[8.4 \times 10^3, 1.3 \times 10^5, -3.6 \times 10^3 ; -7.5 \times 10^3, -3.2 \times 10^3, 9.3 \times 10^3 ; -9.1 \times 10^2, -1.2 \times 10^5, -5.7 \times 10^3\right],\\
\mathrm{K7} &= \left[-1.9 \times 10^3, -3.0 \times 10^5, -1.4 \times 10^4 ; 4.4 \times 10^3, -1.1 \times 10^5, -1.1 \times 10^4 ; -2.4 \times 10^3, 4.0 \times 10^5, 2.5 \times 10^4\right],\\
\mathrm{K8} &= \left[9.2 \times 10^2, 6.5 \times 10^4, 2.4 \times 10^3 ; -2.1 \times 10^3, 1.9 \times 10^4, 3.6 \times 10^3 ; 1.1 \times 10^3, -8.4 \times 10^4, -6.0 \times 10^3\right].
\end{aligned} \tag{40}$$

To test the robustness of the controller, a disturbance

$$W(t) = \left[\sin\frac{(12t)}{3} \quad 0 \quad 0\right]^T \tag{41}$$

is added in (38). For comparison, a controller designed by Theorem 2 of [18] is applied. In [18], the controller was designed for the system with input saturation and disturbance. If the time delay is ignored in [18], its results can be applied in this example. The simulation results are shown in the following figures.

From Figures 3–5, less overshooting can be observed compared with the reference paper. The proposed controller performed well with input saturation, and the H_∞ performance can be achieved.

5. Conclusion

In this paper, the T-S fuzzy control theorem and the ability of contractively invariant ellipsoid were applied for nonlinear systems with input saturation. The proposed method was

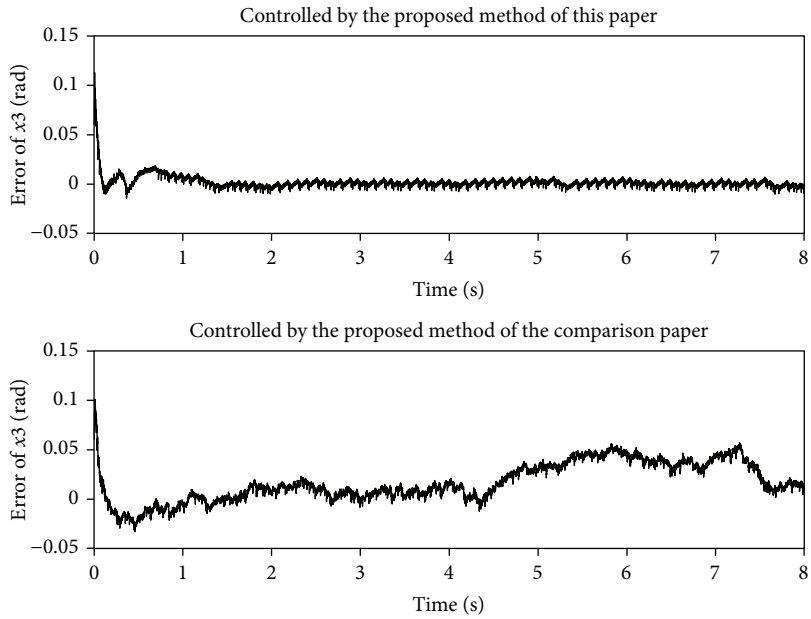


Figure 5: The error of x_3.

described by LMIs, which can be calculated by a computer. The proposed theorem was tested using the Novint Falcon, and a good result was achieved. Comparing the results, the proposed method of this paper has given a better result. Meanwhile, the prescribed H-infinity norm was satisfied. In the example, only eight rules of the T-S fuzzy system were applied; there were still more than one hundred LMIs that need to be calculated. Although a feasible solution of the LMIs can be guaranteed using the proposed method, a method with less LMIs is still necessary for further study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (61601464) and the Fundamental Research Funds for the Central Universities (2013QNB26).

References

[1] T. Hu, A. R. Teel, and L. Zaccarian, "Stability and performance for saturated systems via quadratic and nonquadratic Lyapunov functions," *IEEE Transactions on Automatic Control*, vol. 51, no. 11, pp. 1770–1786, 2006.

[2] S. Sajjadi-Kia and F. Jabbari, "Controllers for linear systems with bounded actuators: slab scheduling and anti-windup," *Automatica*, vol. 49, no. 3, pp. 762–769, 2013.

[3] P. Buckley, "Designing override and feedforward controls," *Control Engineering*, vol. 18, no. 8, pp. 48–51, 1971.

[4] R. Hanus, M. Kinnaert, and J.-L. Henrotte, "Conditioning technique, a general anti-windup and bumpless transfer method," *Automatica*, vol. 23, no. 6, pp. 729–739, 1987.

[5] F.-G. Wang, H.-M. Wang, S.-K. Park, and X.-S. Wang, "Linear pole-placement anti-windup control for input saturation nonlinear system based on Takagi Sugeno fuzzy model," *International Journal of Control, Automation and Systems*, vol. 14, no. 6, pp. 1599–1606, 2016.

[6] V. D. Hajare, A. A. Khandekar, and B. M. Patre, "Discrete sliding mode controller with reaching phase elimination for TITO systems," *ISA Transactions*, vol. 66, pp. 32–45, 2017.

[7] W. Liu, J. Lu, Z. Zhang, and S. Xu, "Observer-based neural control for MIMO pure-feedback non-linear systems with input saturation and disturbances," *IET Control Theory & Applications*, vol. 10, no. 17, pp. 2314–2324, 2016.

[8] A. Lekka, M. C. Turner, and P. P. Menon, "Anti-windup for a class of partially linearisable non-linear systems with application to wave energy converter control," *IET Control Theory & Applications*, vol. 10, no. 18, pp. 2403–2414, 2016.

[9] C. Hua, Y. Yang, and P. X. Liu, "Output-feedback adaptive control of networked teleoperation system with time-varying delay and bounded inputs," *IEEE/ASME Transactions on Mechatronics*, vol. 20, no. 5, pp. 2009–2020, 2015.

[10] L. Zaccarian, Y. Li, E. Weyer, M. Cantoni, and A. R. Teel, "Anti-windup for marginally stable plants and its application to open water channel control systems," *Control Engineering Practice*, vol. 15, no. 2, pp. 261–272, 2007.

[11] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-15, no. 1, pp. 116–132, 1985.

[12] K. Tanaka and H. O. Wang, *Fuzzy Control Systems Design and Analysis*, John Wiley & Sons Inc., New York, NY, USA, 2001.

[13] H. O. Wang, K. Tanaka, and M. F. Griffin, "Parallel distributed compensation of nonlinear systems by Takagi-Sugeno fuzzy model," in *Proceedings of 1995 IEEE International Conference on Fuzzy Systems*, pp. 531–538, Yokohama, Japan, 1995.

[14] J. Willems, "Least squares stationary optimal control and the algebraic Riccati equation," *IEEE Transactions on Automatic Control*, vol. 16, no. 6, pp. 621–634, 1971.

[15] Y. Neestorv and A. Nemirovskii, "A general approach to polynomial-time algorithms design for convex programming," Technical Report, Central Economical and Mathematical Institute, USSR, Moscow, Russia, 1998.

[16] K. Tanaka, T. Ikeda, and H. O. Wang, "Robust stabilization of a class of uncertain nonlinear systems via fuzzy control: quadratic stabilizability, H_∞ control theory, and linear matrix inequalities," *IEEE Transactions on Fuzzy Systems*, vol. 4, no. 1, pp. 1–13, 1996.

[17] S. Xing, Q. Zhang, and B. Zhu, "Mean-square admissibility for stochastic T–S fuzzy singular systems based on extended quadratic Lyapunov function approach," *Fuzzy Sets and Systems*, vol. 307, pp. 99–114, 2017.

[18] S. Xu, G. Sun, and W. Sun, "Takagi–Sugeno fuzzy model based robust dissipative control for uncertain flexible spacecraft with saturated time-delay input," *ISA Transactions*, vol. 66, pp. 105–121, 2017.

[19] W. Guan and F. Liu, "Finite-time dissipative control for singular T–S fuzzy Markovian jump systems under actuator saturation with partly unknown transition rates," *Neurocomputing*, vol. 207, pp. 60–70, 2016.

[20] W. Guan and F. Liu, "Finite-time H_∞ memory state feedback control for uncertain singular TS fuzzy time-delay system under actuator saturation," *Advances in Difference Equations*, vol. 2016, no. 1, 19 pages, 2016.

[21] Y. Ma, P. Yang, and Q. Zhang, "Robust H_∞ control for uncertain singular discrete T-S fuzzy time-delay systems with actuator saturation," *Journal of the Franklin Institute*, vol. 353, no. 13, pp. 3290–3311, 2016.

[22] T. Hu, Z. Lin, and B. M. Chen, "Analysis and design for discrete-time linear systems subject to actuator saturation," *Systems & Control Letters*, vol. 45, no. 2, pp. 97–112, 2002.

[23] W.-J. Chang and Y.-J. Shih, "Fuzzy control of multiplicative noised nonlinear systems subject to actuator saturation and H_∞ performance constraints," *Neurocomputing*, vol. 148, pp. 512–520, 2015.

[24] S. Wen and Y. Yan, "Robust adaptive fuzzy control for a class of uncertain MIMO nonlinear systems with input saturation," *Mathematical Problems in Engineering*, vol. 2015, Article ID 561397, 14 pages, 2015.

[25] A. Nguyen, T. Laurain, R. Palhares, J. Lauber, C. Sentouh, and J. Popieul, "LMI-based control synthesis of constrained Takagi–Sugeno fuzzy systems subject to $\mathscr{L}_2$ or $\mathscr{L}_\infty$ disturbances," *Neurocomputing*, vol. 207, pp. 793–804, 2016.

[26] S. Gao, H. Dong, B. Ning, and X. Yao, "Single-parameter-learning-based fuzzy fault-tolerant output feedback dynamic surface control of constrained-input nonlinear systems," *Information Sciences*, vol. 385-386, pp. 378–394, 2017.

[27] N. Karbasizadeh, A. Aflakiyan, M. Zarei, M. T. Masoulehy, and A. Kalhor, "Dynamic identification of the Novint Falcon haptic device," in *Proceedings of 4-th International Conference on Robotics and Mechatronics*, Tehran, Iran, 2016.

Hindawi
Journal of Sensors
Volume 2018, Article ID 3604520, 8 pages
https://doi.org/10.1155/2018/3604520

Research Article

Performance Analysis of Dual-Polarized Massive MIMO System with Human-Care IoT Devices for Cellular Networks

Jun-Ki Hong

Department of Electrical and Electronic Engineering, Youngsan University, Yangsan, Republic of Korea

Correspondence should be addressed to Jun-Ki Hong; jkhong@ysu.ac.kr

Received 24 November 2017; Accepted 18 February 2018; Published 19 April 2018

Academic Editor: Mucheol Kim

The performance analysis of the dual-polarized massive multiple-input multiple-output (MIMO) system with Internet of things (IoT) devices is studied when outdoor human-care IoT devices are connected to a cellular network via a dual-polarized massive MIMO system. The research background of the performance analysis of dual-polarized massive MIMO system with IoT devices is that recently the data usage of outdoor human-care IoT devices has increased. Therefore, the outdoor human-care IoT devices are necessary to connect with 5G cellular networks which can expect 1000 times higher performance compared with 4G cellular networks. Moreover, in order to guarantee the safety of the patient for emergency cases, a human-care Iot device must be connected to cellular networks which offer more stable communication for outdoors compared to short-range communication technologies such as Wi-Fi, Zigbee, and Bluetooth. To analyze the performance of the dual-polarized massive MIMO system for human-care IoT devices, a dual-polarized MIMO spatial channel model (SCM) is proposed which considers depolarization effect between the dual-polarized transmit-antennas and the receive-antennas. The simulation results show that the performance of the dual-polarized massive MIMO system is improved about 16% to 92% for 20 to 150 IoT devices compared to conventional single-polarized massive MIMO system for identical size of the transmit array.

1. Introduction

According to Business Insider, human-care Internet of things (IoT) devices are expected to grow from 120 million in 2015 to 650 million by 2020 [1]. Furthermore, as the use of wearable human-care IoT devices that attach to the body is increasing, the support of a seamless cellular network is essential instead of short-range communication technologies such as Wi-Fi, Zigbee, and Bluetooth. Since human-care IoT devices are closely related to human health and life, seamless communication is very important to guarantee the safety in case of emergencies especially for outdoors.

In addition, the patient can be monitored in real time by cellular networks for both indoors and outdoors when there are various human-care IoT devices connected to the cellular networks. Especially, the connection between cellular networks and human-care IoT devices provides real-time human-care monitoring system via variable human-care services such as ambulance, smartwatch, first aid kit, and a medically equipped smartphone. Further, the collected information of users can be used to prevent accidents via big data analysis. In contrary, short-range communication technologies such as Wi-Fi, Zigbee, and Bluetooth are difficult to communicate from outdoors. In order to guarantee the safety of the users in case of an emergency, it is necessary to connect with a cellular network which provides more stable connection compared to short-range communication technologies for outdoors.

However, the global data usage is expected to increase to 49 exabytes by 2021 [2] as rapid increase in usage of outdoor human-care-related IoT devices. In recent years, more advanced 5G technology and standards are being actively researched to handle the tremendous amount of data from IoT devices [3, 4].

One of the remarkable 5G technology is the massive multiple-input multiple-output (MIMO) system to increase spectral efficiency by a very large number of transmit-antennas. Typically, the transmit-array of a massive MIMO system is composed of tens or hundreds of transmit-antennas to serve tens or hundreds of users simultaneously

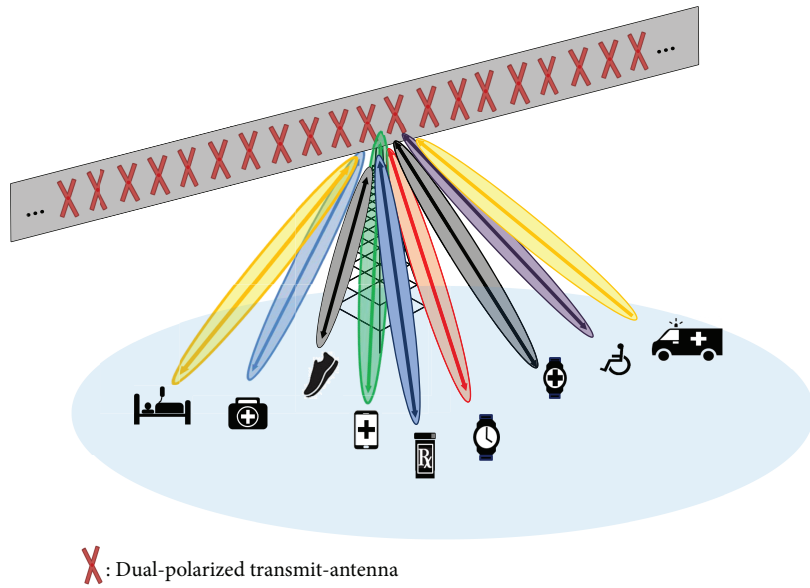


Figure 1: Configuration of the proposed dual-polarized massive MIMO system with various human-care IoT devices for cellular networks.

[5]. Massive MIMO system leads huge improvement in energy efficiency, spectral efficiency robustness, and reliability [6].

A large number of transmit-antenna at a base station (BS) can remarkably improve performance via increasing the multiplexing order to various human-care IoT devices by higher transmit diversity with higher degree of freedoms. In this paper, the performance analysis is achieved when various human-care IoT devices are connected to cellular networks via a dual-polarized massive MIMO system as shown in Figure 1.

To the best of our knowledge, there has not been an analysis that considers the performance between massive the MIMO system and IoT devices for cellular networks. In addition, dual-polarized transmit- and receive-antennas are proposed for BS and IoT device, respectively, to install more antennas and achieve higher performances in limited space.

To evaluate the dual-polarized massive MIMO system, a dual-polarized MIMO spatial channel model (SCM) is proposed which considers the depolarization effects of the dual-polarized antennas. Then, the comparative performance analysis of dual- and single-polarized massive MIMO systems for IoT devices is achieved by system-level simulation.

2. Challenges for Massive MIMO System

Previously, the MIMO system is introduced to increase spectral efficiency by increasing degree of freedoms in space domain by multiple antennas. Afterwards, massive MIMO is introduced to catch up to today's tremendous amount of data traffic with even higher degree of freedoms. However, several issues arise while a large number of transmit-antennas provide more spectral efficiency than conventional MIMO system.

Since transmit-array of massive MIMO is composed of a large number of antennas, the size of the transmit-array will be increased as the number of antenna increases. In other words, the deployment of a large number of transmit-antennas in limited space is very important to implement practical massive MIMO system without unlimited growth of array size.

To deploy a large number of transmit-antennas in limited space, previously, 2 dimensional (2D) planar type of transmitter (Tx) is introduced to reduce the size of the transmit-array [7, 8]. The basic idea of the 2D planar array MIMO system is that the transmit-antennas are installed on a 2D planar grid instead of a 1 dimensional (1D) linear array. Performance comparison between 2D planar and 1D linear transmit-array is analyzed by a function of signal-to-noise ratio (SNR) [7]. 2D planar array performs lower than 1D linear transmit-array because of its higher correlations/interferences between surrounding antennas on the 2D planar array.

Performance analysis of 2D planar and 1D linear transmit-array is conducted by system-level simulations in [7]. Similarly, 2D planar array performs lower than a linear array for same 32 transmit-antennas at the BS. Nevertheless, 2D planar array has several advantages such as reduction of transmit-array size and utilization of vertical beamforming.

3. Dual-Polarized Massive MIMO System

In this paper, a dual-polarized massive MIMO system is proposed which is very effective in reducing the size of a transmit-array by utilization of the extra polarization domain. There are several remarkable advantages for dual-polarized transmit- and receive-antennas since it consists of two colocated orthogonal antennas.

3.1. Advantages of a Dual-Polarized Antenna. One of the major advantages of a dual-polarized transmit-array is the reduction of transmit-array size by half compared to a spatially separate single-polarized transmit-array. Since two colocated orthogonal transmit-antennas are installed on a collocated dual-polarized antenna, a double number of transmit-antennas can be installed compared to a spatially

separate single-polarized transmit-array for identical size of the transmit-array. In other words, the degrees of freedom of a dual-polarized transmit-array can achieve twice the performance compared to a conventional single-polarized transmit-array for identical size of the transmit-array. For example, for a center frequency of 1.9 GHz with a half wavelength, a 64Tx single-polarized transmit-array with a length of 3.67 m can be replaced with 32Tx dual-polarized transmit-array with a length of 1.84 m.

Furthermore, a dual-polarized receive-antenna can be installed at IoT devices to achieve higher performance compared to a single-polarized antenna in identical spaces.

3.2. Utilization of Polarization Domain. Since two independent data streams can be multiplexed at a dual-polarized transmit-antenna, depolarization effect occurs between two data streams. In other words, the extra polarization diversity can be achieved by transmitting two independent data streams from vertically and horizontally polarized antennas at BS. In previous papers, it has been shown that using dual-polarized transmit-antennas can improve performance and reduce transmitter size [9, 10].

However, the performance degradation can occur by polarization mismatch since a single-polarized receive-antenna receives the signal by only one direction and its upper bound of capacity loss from polarization mismatch is 2 bit/s/Hz for multiple dual-polarized antennas and a single-polarized antenna according to [11]. Therefore, the dual-polarized receive-antenna at an IoT device could be the solution to minimize polarization mismatch since vertical and horizontal receive-antenna receives signals at any directions.

Therefore, it is clear that the dual-polarized massive MIMO system is the future for the massive MIMO system to reduce transmit-array size and minimize performance degradation by polarization mismatch. Consequently, the dual-polarized massive MIMO system can be operated by space-polarization division multiple access (SPDMA) which increase spectral efficiency by utilizing both space and polarization domains. The detailed explanations of SPDMA are presented in the next section.

4. SPDMA

The basic idea of SPDMA is increasing spectral efficiency by utilizing both spatial and polarization domains for the dual-polarized MIMO system. In this section, the comparative descriptions of space division multiple access (SDMA), polarization division multiple access (PDMA), and SPDMA techniques are presented.

4.1. SDMA. MIMO system with SDMA is used to extend transmit diversity in the space domain. However, the drawback of the SDMA technique with the single-polarized MIMO system is that the size of a transmit-array increases proportionally as the number of antenna increases. This is one of the major concerns of the massive MIMO system. Moreover, SDMA lacks in utilizing the polarization domain since SDMA only operates for spatially separated single-polarized antennas.

4.2. PDMA. To increase transmit diversity without size expansion, previously, PDMA technique is introduced. The basic idea of PDMA is that the dual-polarized transmit-antenna transmits two independent data streams to different users simultaneously by using orthogonal vertical and horizontal polarizations. The performance analysis between one collocated dual-polarized transmit-antenna and two dual-polarized receive-antennas are provided in [12]. Nevertheless, PDMA lacks spatial diversity at the transmit-array.

4.3. SPDMA. Previous SDMA and PDMA techniques lack polarization and spatial diversity, respectively. In order to overcome these disadvantages, a more advanced SPDMA technique is introduced in [13]. The basic idea of SPDMA is that it increases spectral efficiency by utilizing both space and polarization domains. In other words, the utilization of space domain is achieved by spatially separated dual-polarized antennas and utilization of the polarization domain is achieved by two orthogonal polarizations from vertical and horizontal transmit-antennas. Thus, a one dual-polarized transmit-antenna transmits two different signals simultaneously to different users through the polarization domain. Consequently, the SPDMA technique leads the increment of spectral efficiency not only in space domain but also in polarization domain.

Previously, the comparative performance analysis of the dual- and single-polarized MIMO system is conducted by SDMA and SPDMA techniques via link-level simulation. Simulation results represent the dual-polarized MIMO system with SPDMA achieved higher throughput by extra polarization diversity compared to the single-polarized MIMO with SDMA for identical size of the transmit-array [13]. However, the previous performance analysis of SPDMA is conducted with 8 transmit-antennas which is by a conventional MIMO system. Therefore, previous analysis of the dual-polarized MIMO system with SPDMA is lack of evaluating the performance of a large number of IoT devices.

5. Proposed Dual-Polarized MIMO SCM

3rd Generation Partnership Project (3GPP) MIMO SCM is a geometry-based stochastic model which describes the excess delay, direction of arrival, and direction of departure of multipath for the MIMO channel model [14]. 3GPP MIMO SCM is widely used in modelling the MIMO channels. However, 3GPP MIMO SCM provides a dual-polarized MIMO channel model which considers mixed vertical and horizontal polarizations at a dual-polarized receive-antenna. Thus, previous 3GPP SCM lacks the implementing depolarization effect between vertically and horizontally polarized channels at the dual-polarized receive-antennas which is impossible to consider extra polarization diversity.

Therefore, a dual-polarized MIMO SCM is proposed which considers depolarization effect between dual-polarized transmit- and receive-antennas which operated by the SPDMA technique. Depolarized propagation channels between dual-polarized transmit- and receive-antennas are represented by vertical to vertical (VV), vertical to horizontal

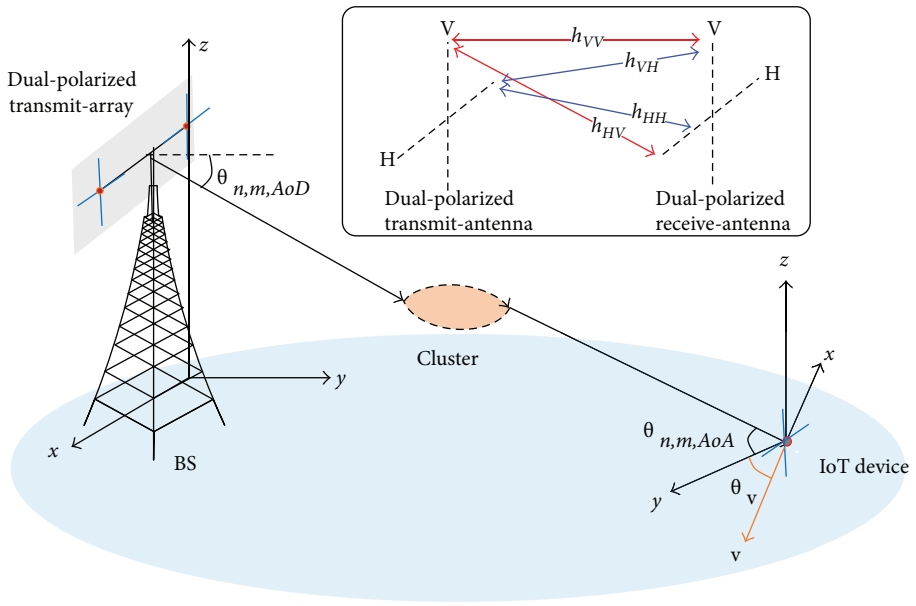


Figure 2: Configuration of the proposed dual-polarized MIMO SCM with an IoT device.

(HV), horizontal to vertical (VH), and horizontal to horizontal (HH) channels as shown in Figure 2.

The proposed dual-polarized MIMO SCM represents BS with S transmit-antenna with S antennas ($S/2 \in N$ of dual-polarized transmit-antennas) and IoT devices with U antennas ($U/2 \in N$ of dual-polarized receive-antennas). Then, (u,s)th dual-polarized MIMO channel component of nth multipath at time t, $\mathbf{H}_{u,s,n}(t)$ can be expressed as

$$\mathbf{H}_{u,s,n}(t) = \sqrt{\frac{P_n \sigma_{SF}}{M}} \sum_{m=1}^{M} \left(\begin{bmatrix} h_{VV} & h_{VH} \\ h_{HV} & h_{HH} \end{bmatrix} \times \exp(\beta_{AoD}) \times \exp(\beta_{\mathrm{AoA}}) \times \exp(\psi_{\mathrm{AoA}}) \right), \tag{1}$$

where

$$\begin{aligned}
h_{VV} &= \chi_{BS,V}(\theta_{n,m,\mathrm{AoD}}) \cdot \left(\sqrt{\frac{\mathrm{XPD}}{1+\mathrm{XPD}}} \exp\left(j\Phi_{n,m}^{(v,v)} \right) \right) \chi_{MS,V}(\theta_{n,m,\mathrm{AoA}}), \\
h_{VH} &= \chi_{BS,H}(\theta_{n,m,\mathrm{AoD}}) \cdot \left(\sqrt{\frac{1}{1+\mathrm{XPD}}} \exp\left(j\Phi_{n,m}^{(v,h)} \right) \right) \chi_{MS,V}(\theta_{n,m,\mathrm{AoA}}), \\
h_{HV} &= \chi_{BS,V}(\theta_{n,m,\mathrm{AoD}}) \cdot \left(\sqrt{\frac{1}{1+\mathrm{XPD}}} \exp\left(j\Phi_{n,m}^{(h,v)} \right) \right) \chi_{MS,H}(\theta_{n,m,\mathrm{AoA}}), \\
h_{HH} &= \chi_{BS,H}(\theta_{n,m,\mathrm{AoD}}) \cdot \left(\sqrt{\frac{\mathrm{XPD}}{1+\mathrm{XPD}}} \exp\left(j\Phi_{n,m}^{(h,h)} \right) \right) \chi_{MS,H}(\theta_{n,m,\mathrm{AoA}}), \\
\beta_{\mathrm{AoD}} &= j2\pi\lambda^{-1}\left\{ d_s \cos(\theta_{n,m,\mathrm{AoD}}) \right\}, \\
\beta_{\mathrm{AoA}} &= j2\pi\lambda^{-1}\left\{ \cos(\theta_{n,m,\mathrm{AoA}}) \right\}, \\
\psi_{\mathrm{AoA}} &= j2\pi\lambda^{-1}\|v\|[\cos(\theta_{n,m,\mathrm{AoA}} - \theta_{\mathbf{v}})].
\end{aligned} \tag{2}$$

The notations are described in Table 1.

Table 1: Notations of the proposed dual-polarized MIMO SCM.

Notation	Description
P_n	The power of the nth path.
σ_{SF}	The lognormal shadow fading.
M	The number of subpaths per path.
$\theta_{n,m,\mathrm{AoD}}$	Absolute angle of departure (AoD) for the mth subpath of the nth path at the BS with respect to the BS broadside.
$\theta_{n,m,\mathrm{AoA}}$	Absolute angle of arrival (AoA) for the mth subpath of the nth pat at the BS with respect to the BS broadside.
$x_{BS,V}(\theta_{n,m,\mathrm{AoD}})$	The BS antenna complex response for the V-pol. Component.
$x_{BS,H}(\theta_{n,m,\mathrm{AoD}})$	The BS antenna complex response for the H-pol. Component.
$x_{MS,V}(\theta_{n,m,\mathrm{AoA}})$	The antenna complex response for the V-pol. Component of the IoT device.
$x_{MS,H}(\theta_{n,m,\mathrm{AoA}})$	The antenna complex response for the H-pol. Component of the IoT device.
$\Phi_{n,m}^{(x,y)}$	The random phase shift between $V(H)$ of the BS and $V(H)$ component of the IoT device for the mth subpath.
$\|\mathbf{v}\|$	The magnitude of the velocity vector of the IoT device.
$\theta_{\mathbf{v}}$	The azimuth angle of the IoT device velocity vector.
j	Square root of -1.
d_u	The distance from BS antenna element s from the reference ($s = 1$) antenna element in meters. For the reference antenna $s = 1$ where $d_1 = 0$.

Aforementioned, the conventional 3GPP MIMO SCM lacks the implementing depolarization effect between dual-polarized transmit- and receive-antennas. Nevertheless, the

Table 2: Simulation parameters.

Parameter	Assumption
Cellular layout	Hexagonal grid, 19 sites, and 3 sectors per site [15]
Simulation scenarios	Urban macro with AS 8° [14]
Sector radius	350 m
Carrier frequency	1.9 GHz
System bandwidth	10 MHz
Channel estimation	Ideal
XPD	0 and 15 dB
Height of BS	35 m
Height of the IoT device	1.5 m
Antenna spacing	0.5λ
BS transmit power	43 dBm
Signal detection algorithm of the IoT device	MMSE
Average speed of moving IoT devices	3 km/h
Noise figure	7 dB
Path loss	COST 231 Hata model [16]

first channel matrix of the proposed channel coefficient (1) represents separated channel propagations between dual-polarized transmit- and receive-antenna by co- and cross-polarizations; VV, HV, VH, and HH with cross-polarization discrimination (XPD) power imbalance terms.

XPD represents the polarization directivity and the discrimination of the power imbalance between copolarization and cross-polarization components at the dual-polarized receive-antenna. Higher XPD value represents less power leakages between vertically and horizontally polarized channels, and lower XPD value represents more power leakages between vertically and horizontally polarized channels in a dual-polarized receive-antenna. A dual-polarized receive-antenna at IoT device is assumed to receive signals from any directions even IoT devices are randomly oriented.

6. Evaluation Method

6.1. TDD Operation. To achieve system-level simulation, time division duplex (TDD) operation is assumed to exploit channel reciprocity in order to simplify channel state information (CSI) estimation for downlink channel from uplink channel. In this case, the training overhead lineally increases as the number of total IoT devices while the number of transmit-antenna can be extended as large as desired.

6.2. System-Level Simulation and Simulation Parameters. To analyze the performance of the dual-polarized massive MIMO system, both dual- and single-polarized massive MIMO systems are evaluated by 3 sector-based system-level simulation which is conducted from [15]. Also, both vertically and horizontally polarized antenna patterns are modeled with 3 dB beam width of 70° with maximum gain of 17 dBi for 3 sector macro cell scenario. Since the goal of this paper is to analyze the performance between proposed dual- and conventional single-polarized massive MIMO systems for identical size of the transmit-array, assuming that there is no performance degradation due to the battery conditions of the IoT devices. The other important parameters are presented in Table 2.

6.3. Dual-Polarized Receive-Antenna with MMSE Receiver at IoT Devices. As mentioned in Section 3.2, single-polarized receive-antenna receives signal by only one direction which leads to performance degradation by polarization mismatch. However, minimum mean square error (MMSE) receiver enables the IoT device to receive either one of the strongest signals or two signals depending on signal-to-interference-plus-noise ratio (SINR) calculated between two colocated orthogonal receive-antennas since IoT devices are composed of a dual-polarized receive-antenna. In other words, a dual-polarized transmit-antenna at BS is able to serve two orthogonal signals to either one IoT device or more than one IoT devices simultaneously depending on the channel conditions to maximize the performances.

6.4. MRT/Multiple IoT Device Selection Algorithm. In this section, the detailed description of the multi-IoT device selection algorithm is presented. It is assumed that the BS simultaneously transmits data to the S IoT devices that have been optimally selected among the K candidate IoT devices The data symbols for the selected S IoT devices are linearly precoded by maximum ratio transmission (MRT), and transmit signal vector is generated as follows:

$$\mathbf{x} = \mathbf{F}\mathbf{b} = \sum_{k=1}^{S} \mathbf{f}_k b_k, \tag{3}$$

where $\mathbf{F} = [\mathbf{f}_1, \mathbf{f}_2 \cdots \mathbf{f}_S]$ is the $S \times S$ precoding matrix where $\mathbf{f}_k$ is the $S \times 1$ precoding vector for selected kth IoT devices among selected S IoT devices. $\mathbf{b} = [b_1, b_2 \cdots b_S]^T$ is the $S \times 1$ transmit symbol vector where b_k is the transmit symbol for the kth IoT device among the selected S IoT devices.

To mitigate the interference between vertically and horizontally polarized receive-antenna at an IoT device, an MMSE receiver is employed based on perfect CSI. Then, the achievable rate of the kth IoT device with an MMSE receiver can be written as

$$C_k = \log_2\left(1 + b_k^* \mathbf{f}_k^H \mathbf{H}_k^H \left(\mathbf{I} + \sum_{l=1, l\neq k}^{M} \mathbf{H}_k \mathbf{f}_l s_l s_l^* \mathbf{f}_l^H \mathbf{H}_k^H\right)^{-1} \mathbf{H}_k \mathbf{f}_k b_k\right), \tag{4}$$

where $\mathbf{H}_k$ and $\mathbf{I}$ denote $U \times S$ dual-polarized MIMO channel of the kth IoT device and $S \times S$ identity matrix, respectively. The precoding vector for MRT to the kth IoT device is denoted by $\mathbf{f}_k$ obtained as

$$\begin{aligned} \mathbf{f}_k &= \mathbf{V}_k, \\ \mathrm{SVD}[\mathbf{H}_k] &= \mathbf{U}_k \mathbf{\Sigma}_k \mathbf{V}_k^H, \end{aligned} \tag{5}$$

where SVD represents the singular value decomposition. $\mathbf{U}_k$, $\mathbf{\Sigma}_k$, and $\mathbf{V}_k$ represent $S \times S$ complex unitary matrix, $S \times U$

initialize $G_{opt.} = \{\varnothing\}, C_{\max} = 0, a = 0$
do
$n = \arg\max_{n \in O, n \neq G_{opt.}} \sum_{k \in \{n, G_{opt.}\}} \log_2(1 + \gamma_k)$
$C = \sum_{k \in \{n, G_{opt.}\}} \log_2(1 + \gamma_k)$
if $C > C_{\max.}$.
$G_{opt.} = \{G_{opt.}, n\}$
$R_{\max} = R$
$a = 1$
end
while $a = 1$ **and** $|G_{opt.}| < S_{tx}$

Algorithm 1: Proposed IoT device selection algorithm.

rectangular diagonal matrix, and $U \times U$ complex unitary matrix of the kth IoT device, respectively.

To evaluate the throughput of the proposed massive MIMO systems, the multiuser scheduling algorithm is proposed which serves multiple IoT devices simultaneously with highest IoT-sum-rate by MRT precoding at BS. By MRT, the signals are transmitted by the strongest eigenmodes at BS and the received signals at an IoT device are combined by maximal ratio combining (MRC) technique. The proposed IoT device selection algorithm is specified as follows.where O denotes the universal set of IoTs. S_{tx}, C, and $C_{\max}$ represent the number of transmit-antennas, throughput, and the maximum throughput of the kth IoT device, respectively.

7. Performance Evaluation

7.1. Simulation Scenarios and Level of XPD. To analyze the performance of the proposed dual-polarized MIMO system, three simulation scenarios are proposed: 32Tx dual-polarized massive MIMO system, 32Tx single-polarized massive MIMO system, and 8Tx single-polarized MIMO system. 32Tx single-polarized massive MIMO system represents the identical size of transmit-array of the proposed 32Tx dual-polarized massive MIMO system. Further, 8Tx single-polarized MIMO system represents the reference of conventional MIMO system to compare performance with proposed massive MIMO systems. Moreover, IoT devices are assumed as one dual-polarized receive-antenna to receive signals by any directions.

To investigate the impact of XPD variations on the proposed simulation scenarios, XPD values of 0 and 15 dB are implemented while the proposed dual-polarized MIMO SCM consists XPD terms. XPD = 15 dB is presumed for high level of XPD according to simulation result in [17].

7.2. Performance Analysis. In this subsection, simulation results of the system-level simulation are carried out to analyze the performance of the dual-polarized massive MIMO system compared to single-polarized massive MIMO systems. As mentioned in Subsection 6.4, the performances of the proposed MIMO systems are evaluated by highest sum-rate capacity from the multi-IoT scheduling.

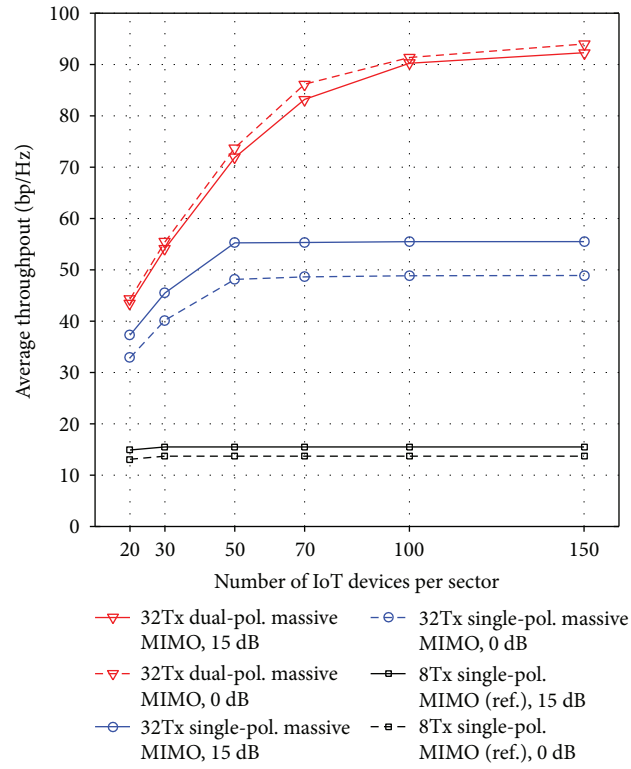


Figure 3: Comparison of the average throughputs.

Figure 3 represents the comparison of the average throughputs between the dual- and single-polarized massive MIMO systems as a function of the number of IoT devices which represents the spectral efficiency. As shown in Figure 3, a proposed 32Tx dual-polarized massive MIMO continuously increases the throughput over 150 IoT devices while a conventional 32Tx single-polarization massive MIMO is impossible to increase the throughputs when more than 50 IoT devices are present in the sector. The performance of the 32Tx dual-polarized massive MIMO system has improved by 16% to 92% for 20 to 150 IoT devices compared to a conventional 32Tx single-polarized massive MIMO system for identical size of the transmit-array.

Furthermore, the average throughput of the dual-polarized massive MIMO for 0 dB is higher than that of the 15 dB because a dual-polarized receive-antenna prefers to receive uncorrelated signals simultaneously instead of selecting one of the strongest signals between two transmitted signals. In other words, completely uncorrelated received signals benefit to increase average sum for the dual-polarized massive MIMO system. Furthermore, the colocated two orthogonal receive-antennas have benefited from uncorrelated channels since less coupling and interferences are generated between two orthogonal receive-antennas. Moreover, the dual-polarized massive MIMO system is insensitive to XPD variations according to throughput gap between XPD = 0 dB and XPD = 15 dB compared to the performance of the single-polarized massive MIMO system.

On the other hand, the single-polarized massive MIMO systems perform higher average throughput for XPD = 15 dB since a higher level of XPD represents the higher directivity

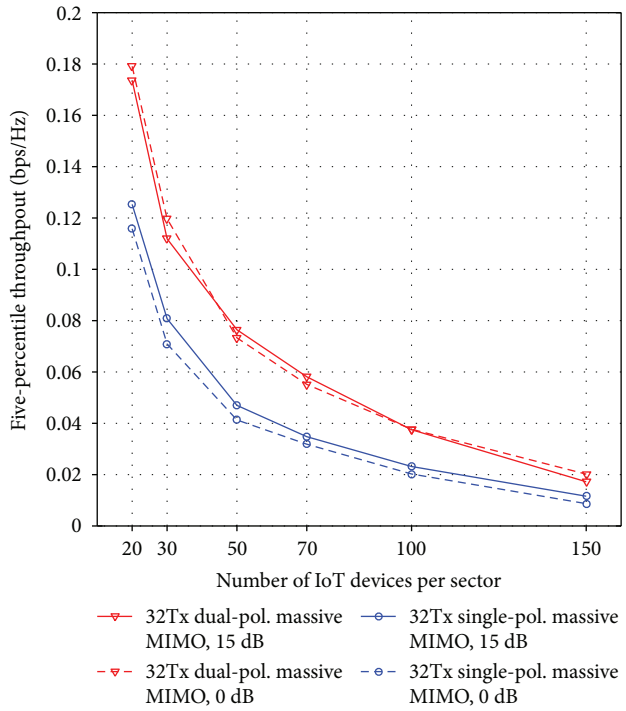


Figure 4: Five-percentile throughputs of the dual- and single-polarized massive MIMO systems.

between transmit and receive antennas. Therefore, transmit-antennas of single-polarized massive MIMO systems have benefited from high XPD since the coupling effect and interference are not generated from colocated transmit-antenna. Since XPD = 0 dB represents the general environment, it is a disadvantage for the single-polarized massive MIMO system.

Further, a conventional 8Tx single-polarized MIMO system is impossible in increasing the performance as the number of IoT devices increases, but the proposed 32Tx dual-polarized massive MIMO system continuously increases the throughput by a higher degree of freedom with the transmit and polarization diversity.

Figure 4 presents the five-percentile (or sector-edge) throughput of the dual- and single-polarized massive MIMO systems. These simulation results indicate that the proposed dual-polarized massive MIMO improves the performance compared to 32Tx single-polarized massive MIMO systems by identical size of the transmit-array.

According to simulation results from Figures 3 and 4, the performance of the dual-polarized massive MIMO system is improved compared to the single-polarized massive MIMO system for identical size of the transmit-array since two colocated polarized transmit-antennas aid to serve more IoT devices via the polarization domain. Consequently, the dual-polarized massive MIMO system benefits from uncorrelated channel environment which is very useful at XPD = 0 dB which represents the general environment.

8. Conclusion

In this paper, the performance between human-care IoT devices and the massive MIMO system is analyzed while massive MIMO system provides seamless communication and big data processing especially for outdoor human-care IoT devices. According to simulation results, the proposed dual-polarized massive MIMO system has several advantages such as reduction of transmit-array size and utilization of polarization domain of the transmit-array. Simulation results show that the performance of the dual-polarized massive MIMO is improved for uncorrelated channel environment for outdoor, XPD = 0 dB. Furthermore, the proposed dual-polarized massive MIMO system doubles the performance compared to the conventional single-polarized massive MIMO system via the identical size of the transmit-array for a large number of IoT devices.

Thus, the proposed dual-polarized massive MIMO system can be the solution to improve performances for outdoor human-care IoT devices while dual-polarized massive MIMO systems enable in reducing the sizes of the transmitters and receivers.

Conflicts of Interest

The author declares that he has no conflicts of interest.

References

[1] Business Insider, *Internet of Things in Healthcare: Information Technology in Health*, December 2016, http://www.businessinsider.com/internet-of-things-in-healthcare-2016-8.

[2] Cisco, *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2014–2019*, http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.pdf.

[3] M. R. Palattella, M. Dohler, A. Grieco et al., "Internet of things in the 5G era: enablers, architecture, and business models," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 3, pp. 510–527, 2016.

[4] G. A. Akpakwu, B. J. Silva, G. P. Hancke, and A. M. Abu-Mahfouz, "A survey on 5G networks for the internet of things: communication technologies and challenges," *IEEE Access*, vol. PP, no. 99, p. 1, 2017.

[5] T. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, 2010.

[6] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, 2014.

[7] J.-K. Hong, H. Joung, H.-S. Jo, C. Mun, and J.-G. Yook, "A three dimensional model for dual polarized MIMO channel of the planar array antennas," *International Journal of Smart Home*, vol. 7, no. 1, 2013.

[8] Y.-H. Nam, B. Ng, K. Sayana et al., "Full-dimension MIMO (FD-MIMO) for next generation cellular technology," *IEEE Communications Magazine*, vol. 51, no. 6, pp. 172–179, 2013.

[9] J. Jootar, J.-F. Diouris, and J. R. Zeidler, "Performance of polarization diversity in correlated Nakagami-m fading channels," *IEEE Transactions on Vehicular Technology*, vol. 55, no. 1, pp. 128–136, 2006.

[10] R. U. Nabar, H. Bolcskei, V. Erceg, D. Gesbert, and A. J. Paulraj, "Performance of multiantenna signaling techniques in the presence of polarization diversity," *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2553–2562, 2002.

[11] H. Joung, H.-S. Jo, C. Mun, and J.-G. Yook, "Capacity loss due to polarization-mismatch and space-correlation on MISO channel," *IEEE Transactions on Wireless Communications*, vol. 13, no. 4, pp. 2124–2136, 2014.

[12] S.-C. Kwon and G. L. Stuber, "Polarization division multiple access on NLoS wide-band wireless fading channels," *IEEE Transactions on Wireless Communications*, vol. 13, no. 7, pp. 3726–3737, 2014.

[13] H. Joung, H.-S. Jo, C. Mun, and J.-G. Yook, "Space-polarization division multiple access system with limited feedback," *KSII Transactions on Internet and Information Systems*, vol. 8, no. 4, pp. 1292–1306, 2014.

[14] 3GPP TR 25.996 V14.0.0, "Spatial channel model for multiple input multiple output (MIMO) simulations," Tech. Rep., 2017, http://www.etsi.org/deliver/etsi_tr/125900_125999/125996/14.00.00_60/tr_125996v140000p.pdf.

[15] R. Srinivasan, Ed., *IEEE 802.16m Evaluation Methodology Document (EMD)*, 2008, http://ieee802.org/16/tgm/docs/80216m-08_004r2.pdf.

[16] E. Damosso, L. M. Correia, and European Commission, "COST Action 231: Digital mobile radio towards future generation systems: final report," 1999, http://www.lx.it.pt/cost231/final_report.htm.

[17] J.-U. Jeong, J.-H. Kim, and C. Mun, "Analysis of massive MIMO wireless channel characteristics," *The Journal of Korea Information and Communications Society*, vol. 38B, no. 3, pp. 216–221, 2013.

Hindawi
Journal of Sensors
Volume 2018, Article ID 7168524, 10 pages
https://doi.org/10.1155/2018/7168524

Research Article

A Development of Clinical Decision Support System for Video Head Impulse Test Based on Fuzzy Inference System

Dao Thi Anh Nguyen,[1] Insu Won,[2] Kyusung Kim,[3] and Jangwoo Kwon[1]

[1]*Department of Applied Physics, University of Eastern Finland, 70211 Kuopio, Finland*
[2]*Institute of Electronics and Information Research, Inha University, Incheon 22212, Republic of Korea*
[3]*Department of Otorhinolaryngology, College of Medicine, Inha University, Incheon 22332, Republic of Korea*

Correspondence should be addressed to Jangwoo Kwon; jwkwon@inha.ac.kr

Received 15 November 2017; Accepted 5 February 2018; Published 2 April 2018

Academic Editor: Ka L. Man

This paper represents the clinical decision support system for video head impulse test (vHIT) based on fuzzy inference system. It examines the eye and head movement recorded by the eye movement tracking device, calculates the vestibulo-ocular reflex (VOR) gain, and applies fuzzy inference system to output the normality and artifact index of the test result. The position VOR gain and the proportion of covert and overt catch-up saccades (CUS) within the dataset are used as the input of the inference system. In addition, this system yields one more factor, the artifact index, which represents the current interference in the dataset. Data of fifteen vestibular neuritis patients and two of normal subjects were evaluated. The artifact index appears to be very high in the lesion side of vestibular neuritis (VN) patients, indicating highly theoretical contradictions, which are low gain but without CUS, or normal gain with the appearance of CUS. Both intact side and normal subject show high normality and low artifact index, even though the intact side has slightly lower normality and higher artifact index. In conclusion, this is a robust system, which is the first one that takes gain and CUS into account, to output not only the normality of the vHIT dataset, but also the artifacts.

1. Introduction

The vestibulo-ocular reflex (VOR) is a dynamic vestibular function, which helps humans to maintain balance and clear vision during head rotations or translations. This reflex, which depends on the vestibule, the acceleration detectors of the inner ear, generates eye movements at short latency (<15 ms) to compensate for head movement (rotations or displacements—translations) [1]. The VOR gain, the compensation ratio of the eye movement to head movement, accordingly, is the decisive factor to measure this dynamic vestibular function. However, there are many systems, apart from the semicircular canals, controlling eye movements [1]. Voluntary saccadic, smooth pursuit, visual (optokinetic) input, or cervical input, can all control eye movements, and in order to test semicircular canal function specifically, the contribution of these additional sources of control must be excluded [2]. Fortunately, only the VOR reflexes appear in high-frequency movement. Hence, in order to examine the patient's function at particularly high acceleration, the head impulse test was first described by Cremer et al. and Weber et al. and became a specialized clinical assessment for VOR function. This test examines the function of semicircular canals in the inner ear at the high angular acceleration ($2000^\circ/s^2$–$4000^\circ/s^2$) with the narrow angular extent of 5°–15° [3–9]. In order to perform this test properly, scleral search coil technique was innovated. However, this system is bulky and not practical, especially in frontline healthcare. That is the reason why it has recently been replaced by the portable video head impulse test (vHIT) which enables this test to examine all six semicircular canals in both ears and detect the eye movements that are imperceptible to the naked eye. However, the practical assessment for vHIT still confronts with lots of artifacts; 72% had abnormal disruptive saccades, 44% had at least one artifact, and 42% were uninterpretable. These might come from the recording device (goggle slippage, improper pupil tracking algorithm, etc.), from the patient cases (eye disorders, kids, etc.), and

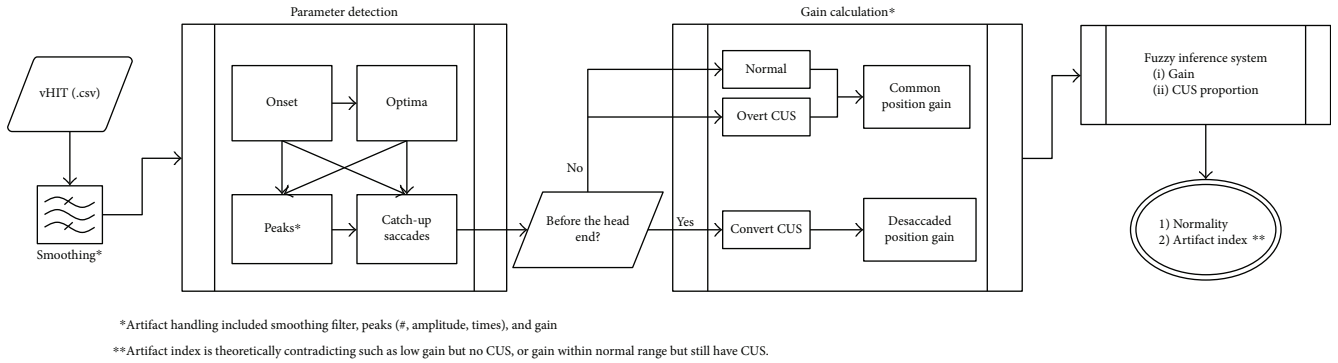


Figure 1: System block diagram. After loading the data file (.csv), the head impulse and eye response are first smoothed, then used for detecting parameters. Those include onset, minimum and maximum values, significant peaks, and catch-up saccades (CUS). If the CUS appears before the head impulse end, it is labeled as covert CUS, and the gain is calculated via desaccaded position gain algorithm. Otherwise, it can be a normal response, or contains overt CUS, which can be separated due to the number and time of significant peaks. After all, the gain and CUS proportion is calculated and inputted to fuzzy inference system to obtain normality and artifact index. *Artifact handling included smoothing filter, peaks (#, amplitude, and times), and gain. **Artifact index is theoretically contradicting such as low gain but no CUS, or gain within normal range but still have CUS.

the examining environment (lighting). As a result, expert assessment and interpretation of vHIT results are still required [10]. To overcome this drawback, this research proposes a clinical decision support system specialized in interpretation of video head impulse test's examination result using fuzzy logic. A clinical decision support system is a health information technology system that is designed to provide physicians and other health professionals with clinical decision support, that is, assistance with clinical decision-making tasks [11, 12]. Furthermore, fuzzy logic was applied in this research due to its compliance with the partial truth, in which the true value might range between fully truthful or false. Instead of defining one specific value into a bivalent if-then condition, a membership function for each subrange of each variable was employed, which describes how much one value partially belongs to one subrange. As a result, the fuzzy set extends the classical set, extending the fuzzy inference system to multivalued logic and the conventional if-then condition to linguistic variables, making it close to the natural languages and human reasoning. That is the reason why we applied the fuzzy logic theory into clinical decision support system, which can resemble the clinical experts in vHIT assessment.

Two important factors in vHIT are the normality and the artifact of the dataset. The normality is defined based on the VOR gain and catch-up saccades (CUS) presented in that dataset. The VOR gain and CUS of the impulse, however, are feeble toward artifacts so it is sometimes not reliable enough to determine the VOR purely based on the VOR gain and the appearance of CUS, especially without reviewing individual impulses because one noisy impulse can lead to wrongly detect the CUS in the whole set. Hence, this system combines both VOR gain and the proportion of CUS to yield the normality and the artifact index using one knowledge-based clinical decision support system. This system is an upgraded version for the previous fuzzy logic-based recommendation system [13]. In [13], two types of normality were employed based on gain and/or covert CUS and obtained by the two inference systems. The first system uses only gain as an input, whereas the second one uses both gain and covert CUS appearance. This system, however, is naïve to the artifact and implements only covert CUS appearance which is not enough to define the normality of the dataset. In our new system, the inputs were updated to the proportion of both covert and overt CUS and VOR gain. Furthermore, it outputs the artifact index, which represents how conflicting it is theoretically. The new fuzzy inference system's block diagram is shown in Figure 1.

2. Methods

2.1. Catch-Up Saccade Detection. During the head impulse test, if VOR eye movements fail to keep the eyes on target, the difference between the head and eye positions triggers some adaptive motor responses, providing additional eye compensatory pursuit of eye to head movement, so-called CUS. These corrective saccadic movements might occur after head impulses, resulting in overt catch-up saccades, which are visible and can be detected by an experienced examiner during the bedside test without any additional equipment. If the position difference between the head and eye is predicted early enough, resulting in the short saccadic latency, catch-up saccades can happen during head impulses, become invisible to the clinician's naked eye, resulting in covert catch-up saccades. These kinds of eye movements are practically impossible to detect without specialized equipment and are not likely to reposition the eyes exactly on the target when the head impulses are unpredictable. The latency of catch-up saccades (100 ms) is considerably more than the latency of VOR eye movements (15 ms) because of the cortical involvement [1]. The CUS was statistically found with three conditions of refixation saccades occurred frequently in cases with abnormal hVOR: isolated covert saccades (13.7%), isolated overt saccades (34.3%), and the combination of overt and covert saccades (52.0%) [14]. Another motor response, which appears in vHIT eye velocity, is spontaneous

nystagmus. This is a rapid involuntary movement of the eyes and is mainly caused by a central lesion [15]. In this study, a spontaneous nystagmus was ignored and grouped with overt catch-up saccades for further examination.

One problem is that the covert CUS occurs during the head impulse. Consequently, it strongly influences on the gain calculation.

In this research, one catch-up saccade is categorized into covert set if its peak occurs from head peak until the zero crossing of head movement. On the other hand, overt catch-up saccades appear after the head impulse. So, if the saccades' peak is detected after zero crossing of the head movement, but not later than 450 ms after the head movement onset, that peak is defined as overt catch-up saccade. The peaks that appear later than 450 ms, however, are neglected.

2.2. VOR Gain Calculation. The VOR gain represents the compensation of the eye movement to the head movement. Cremer et al. have reported that unilateral vestibular lesions result in a permanent reduction in the VOR gain during sudden head thrusts toward the side of the lesion [3]. In normal cases, the VOR gain is close to one, then gradually decreases when the head velocity peak is higher than 200°/s. In unilateral vestibular deficit cases, the VOR gain is significantly below one for head impulses toward the side of lesion and declines rapidly with increasing head velocity. In the intact side, the VOR gain is similar but not as much as in the lesion side. In bilateral vestibular deficit, the VOR gain toward both sides is significantly less than one [6].

The VOR gain calculated based on the head and eye movements can be defined based on velocity, acceleration, or position. At first, velocity VOR gain is widely used and calculated by several approaches. As in [6], horizontal velocity VOR gain was calculated as the ratio of mean eye velocity over mean head velocity during a 40 msec window centered at the peak head acceleration. Or else, velocity VOR gain in [16] was calculated for each subject by dividing the length of the total eye velocity vector (eye speed) by the length of the head velocity vector (head speed). The disadvantage of the device that uses velocity VOR gain as patient output is that it requires a lot of careful calibration. Moreover, the recorded velocity is easier to be affected by noise (from device, clinician, and subject) as compared to position VOR gain [17]. Consequently, the position VOR gain is used in this research. Its mechanism is that firstly the head movement is the head velocity cumulated from the head onset until the next zero crossing of head velocity trace. Then, the eye movement is the eye velocity cumulated from the eye onset until the next zero crossing of the eye velocity trace. Finally, the VOR gain is the ratio of the eye movement to the head movement. This calculation is different from ICS device, which uses the frame of calculating gain between the head onset until its zero crossing [6]. This might be unstable due to the eye response's timing, which is theoretically 15 ms after the head movement [1]. In our dataset, the eye response might occur earlier or has bigger latency.

In addition, the position VOR gain calculation was optimized by removing covert CUS, because this type of CUS strongly influences on gain calculation. This method is called desaccaded position gain [7]. In this study, after a covert CUS is detected, a line drawn from the valley of covert CUS to the next zero crossing of eye movement is used to remove the covert CUS. Afterwards, the position VOR gain is calculated as mentioned above.

2.3. Artifact Handling

2.3.1. Smoothing Filter. The velocity traces of both the head and eye contain small ripples, which possibly influence on the gain calculation and saccade detection. Hence, a smoothing filter is recommended. We choose Savitzky-Golay smoothing filter (SGF), a digital filter which minimizes the least squares error in fitting a low-degree polynomial to partial windows of a noisy data. Furthermore, it performs better than standard averaging finite impulse response filters, because the velocity traces do not contain frequent noise, but white noise. The disadvantage of SGF is that it is less successful than the conventional one at rejecting noise. However, in our application, signal preservation has prior consideration. In this research, SGF with the third order and window size of 13 was found as an optima parameter for head velocity trace [18].

2.3.2. vHIT's Artifact. One problem that might appear is that the plot will look extremely noisy if there are so many impulses, or one noisy impulse contains more than three high-eye velocity peaks. This could lead to the clinician confusion with the CUS, resulting in wrong detection. Then the proportion should be taken into account.

As the counting method, the combination of gain and proportion of catch-up saccades contains two cases contradicting with the theory when assessing the vHIT dataset. One is the response's gain is within normal or high range, but there is CUS. This might be the goggle slippage and so on. The second one is the patient has low gain but no CUS. This has an unknown reason.

The fuzzy inference system does come with an output of abnormality, in which can be defined into two indices: known reason and unknown reason.

2.4. Patients. Eight different types of vHIT's artifacts: phase shift, inappropriately high gain, pseudo-saccades, multiple VOR peaks, excessive post-HIT bounce, eye moves opposite the expected slow phase VOR direction, trace oscillations (noisy baseline), and unclassifiable artifacts (i.e., multiple different artifact morphologies or unrecognizable morphologies that were clearly nonphysiologic). These artifacts might come from patient factor, human error, detection algorithm, or unknown artifacts [10]. With an assumption, the normality of the response will decrease with the increase of the proportion.

The datasets of fifteen vestibular neuritis (VN) subjects and two normal controls were used to evaluate this system. All fifteen patients had acute unilateral vestibular neuritis diagnosed from clinical symptoms and a vestibular function test (mean 46 years, age range 22–69, and one female). The first twelve VN subjects have vestibular deficit on the left side; the last three have on the right side.

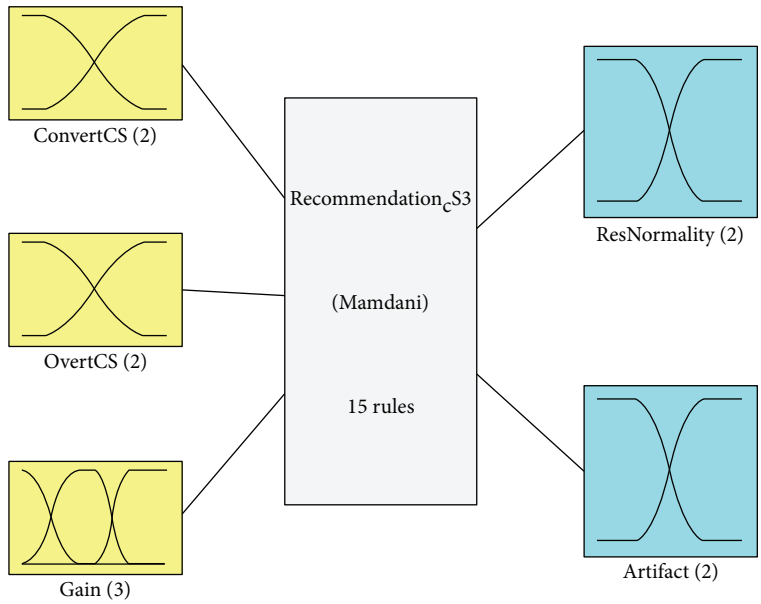


Figure 2: Fuzzy inference system with three inputs: covert CUS proportion, overt CUS proportion, and gain; two outputs, normality and artifact. This system has a rule set of 15 weighted rules, which are built based on personal knowledge.

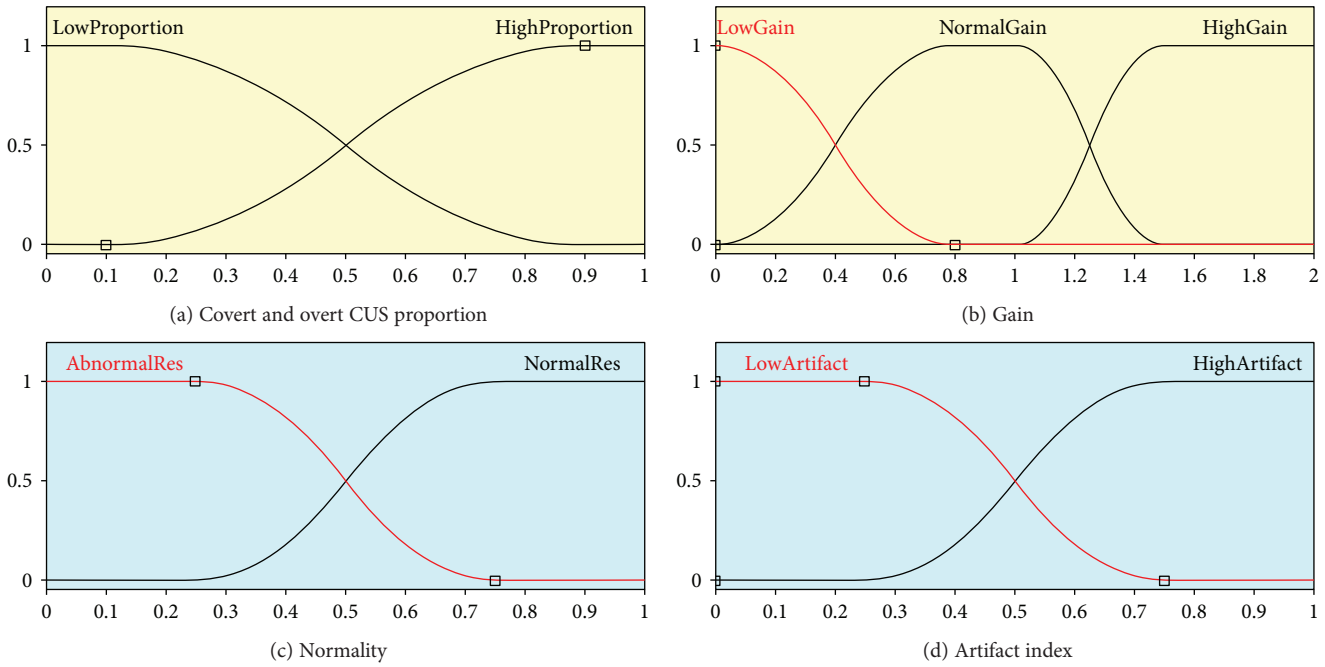
2.5. Fuzzy Logic Inference System. Fuzzy logic is an extension of Boolean logic based on the mathematical theory of fuzzy sets. By introducing fuzziness, called the membership function, thus enabling a conditional state other than just true or false, fuzzy logic provides valuable and flexible reasoning, which makes it possible to take into account inaccuracies and uncertainties. One advantage of using fuzzy logic to formalize human reasoning is that its rules are set in natural language. That is why fuzzy logic is applicable for constructing a human-readable collaborative recommendation application [19, 20]. In the conventional Boolean logic, for instance, a gain of 0.6 is categorized into abnormal set, but it is not truly deficient in defining the gain. On the other way, the fuzzy set gives a degree of truth for that exact value, for example, 0.6 is abnormal, but 0.4 is normal.

Fuzzy inference is a method that interprets the values in the input vector and, based on some set of rules, assigns values to the output vector. Fuzzy inference system helps to bring out the input-output mapping using fuzzy logic. In order to do so, its structure consists of three main parts: membership function, logical operation, and if-then rules. The structure of fuzzy inference system consists of five parts: fuzzification of the input variables, application of the fuzzy operator (AND or OR) in the antecedent, implication from the antecedent to the consequent, aggregation of the consequents across the rules, and finally defuzzification. All of these steps will be described below [21, 22].

In this CDSS, we have five membership functions: three inputs (CovertCS, OvertCS, and Gain) and two outputs (ResNormality, Artifact) as shown in Figure 2. Velocity VOR gain was used in this reference system with the normal range: 0.8–1.2. The normal range of position VOR gain in this inference system is 0.8-1.2, which is higher than that of velocity VOR gain. This is because position VOR gain contains variance due to the fluctuation during the whole period of head impulse. Overt CUS is detected as the CUS after the head stop, but not later than 550 ms. Covert CUS is the CUS occurring during the head impulse. Both the CUS proportions were partially defined by Z-shaped membership function as in 1 (LowProportion; 0.1, 0.9) and S-shaped membership function as in 2 (HighProportion; 0.1, 0.9), as shown in Figure 3(a). The gain membership function is partially divided into three π-shaped functions defined by 3: (LowGain; 0, 0, 0, and 0.8), (NormalGain; 0, 0.8, 1, and 1.5), and (HighGain; 1, 1.5, 2, and 2), as shown in Figure 3(b). The membership functions for normality (ResNormality) have two π-shaped functions (AbnormalRes; 0, 0, 0.25, and 0.75) and (NormalRes; 0.25, 0.75, and 1, 1), as shown in Figure 3(c). The membership functions for artifact index (artifact) was defined by two π-shaped functions (LowArtifact; 0, 0, 0.25, and 0.75) and (HighArtifact; 0.25, 0.75, and 1, 1) as shown in Figure 3(d). Consequently, once the crisp input is presented, it will be fuzzificated into a set of values for each of its membership functions.

(i) Z-shaped membership function:

$$f_{\text{Z-shaped}}(x; a, b) = \begin{cases} 1, & x \le a, \\ 1 - 2\left(\dfrac{x-a}{b-a}\right)^2, & a < x \le \dfrac{a+b}{2}, \\ 2\left(\dfrac{x-b}{b-a}\right)^2, & \dfrac{a+b}{2} < x \le b, \\ 0, & x > b. \end{cases} \quad (1)$$



(a) Covert and overt CUS proportion

(b) Gain

(c) Normality

(d) Artifact index

Figure 3: Membership function for each input and output.

(ii) S-shaped membership function:

$$f_{\text{S-shaped}}(x;a,b)=\begin{cases}0, & x\le a,\\ 2\left(\dfrac{x-a}{b-a}\right)^2, & a<x\le\dfrac{a+b}{2},\\ 1-2\left(\dfrac{x-b}{b-a}\right)^2, & \dfrac{a+b}{2}<x\le b,\\ 1, & x>b.\end{cases}\tag{2}$$

(iii) π-shaped membership function:

$$f_{\pi\text{-shaped}}(x;a,b,c,d)=\begin{cases}0, & x\le a,\\ 2\left(\dfrac{x-a}{b-a}\right)^2, & a<x\le\dfrac{a+b}{2},\\ 1-2\left(\dfrac{x-b}{b-a}\right)^2, & \dfrac{a+b}{2}<x\le b,\\ 1, & b<x\le c,\\ 1-2\left(\dfrac{x-c}{d-c}\right)^2, & c<x\le\dfrac{c+d}{2},\\ 2\left(\dfrac{x-d}{d-c}\right)^2, & \dfrac{c+d}{2}<x\le d,\\ 0, & x>d.\end{cases}\tag{3}$$

After the membership function was specified, the set of 15 if-then rules for three inputs and two outputs was independently defined as shown in Figure 4. These rule sets represent the relationship among each input and output using linguistic values of each parameter. The antecedent, if statement of inputs, combines the value set of input which was obtained from membership function using fuzzy operator. In this CDSS, we use AND operator. Afterwards, these results will be implicated using the rule's weight. The implication method in this research is min operator. Then the result is summed up, so-called the aggregation, and finally, the outputs are defuzzified into specified value. Based on different implication relations, different outputs were derived. In this research, we used Mamdani-type inference system because of its intuitiveness as human input and its wide acceptance. The input-output mapping was represented in Figure 5.

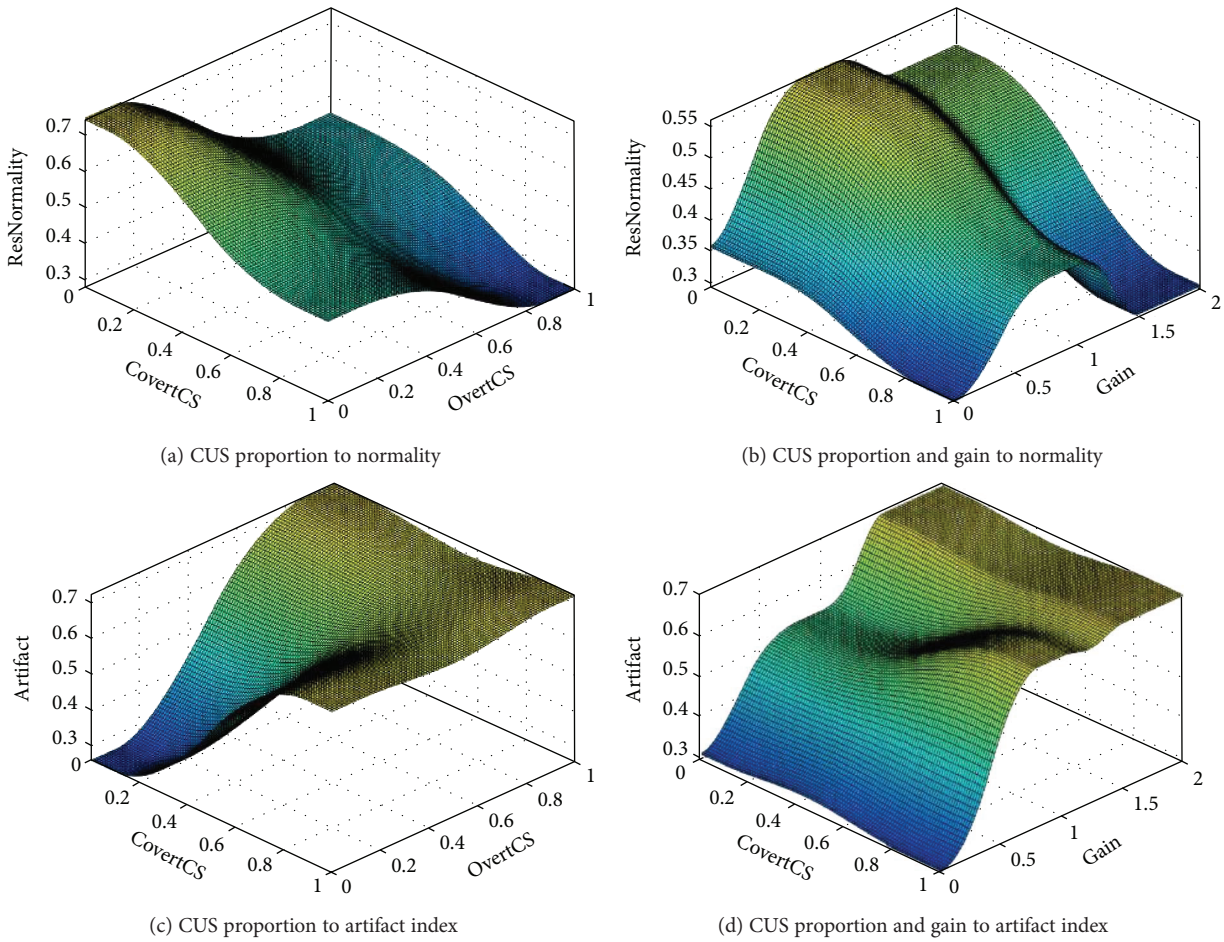
Finally, yet importantly, the membership function, rule set, and rule's weight were intuitively defined based on clinical assessment and under supervision of the experts. Moreover, with sufficient data, this system can be converted to Sugeno-type fuzzy inference, which can be upgraded to the adaptive neural inference system in order to implicitly learn from the database.

3. Results

3.1. Fuzzy Inference System. The new application consists of two recommendation applications. In the previous systems, the fuzzy inference system was used to convert position gain and appearance of covert catch-up saccades (from experiment 1) into the efficiency of the subject's vestibule function, which is more significant to the application user. In contrast, in the new system, these two applications are a Mamdani type, where the implementation method is minimum, the aggregate method is maximum, and defuzzification is the minimum of the maximum. The first is Recommendation 1, one input for which is position gain. The

1. If (CovertCS is LowProportion) and (OvertCS is LowProportion) and (Gain is LowGain) then (ResNormality is NormalRes)(Artifact is LowArtifact) (0.2)
2. If (CovertCS is LowProportion) and (OvertCSis HighProportion) and (Gain is LowGain) then (ResNormality is AbnormalRes)(Artifact is LowArtifact) (1)
3. If (CovertCS is HighProportion) and (OvertCS is LowProportion) and (Gain is LowGain) then (ResNormality is AbnormalRes)(Artifact is LowArtifact) (1)
4. If (CovertCS is HighProportion) and (OvertCS is HighProportion) and (Gain is LowGain) then (ResNormality is AbnormalRes)(Artifact is LowArtifact) (1)
5. If (CovertCS is LowProportion) and (OvertCS is LowProportion) and (Gain is NormalGain) then (ResNormality is NormalRes)(Artifact is LowArtifact) (1)
6. If (CovertCS is LowProportion) and (OvertCS is HighProportion) and (Gain is NormalGain) then (ResNormality is AbnormalRes)(Artifact is HighArtifact) (0.8)
7. If (CovertCS is HighProportion) and (OvertCS is LowProportion) and (Gain is NormalGain) then (ResNormality is AbnormalRes)(Artifact is HighArtifact) (0.5)
8. If (CovertCS is HighProportion) and (OvertCS is HighProportion) and (Gain is NormalGain) then (ResNormality is AbnormalRes)(Artifact is HighArtifact) (0.7)
9. If (CovertCS is LowProportion) and (OvertCS is LowProportion) and (Gain is HighGain) then (ResNormality is NormalRes)(Artifact is HighArtifact) (0.7)
10. If (CovertCS is LowProportion) and (OvertCS is HighProportion) and (Gain is HighGain) then (ResNormality is AbnormalRes)(Artifact is HighArtifact) (0.7)
11. If (CovertCS is HighProportion) and (OvertCS is LowProportion) and (Gain is HighGain) then (ResNormality is AbnormalRes)(Artifact is HighArtifact) (0.7)
12.If (CovertCS is HighProportion) and (OvertCS is HighProportion) and (Gain is HighGain) then (ResNormality is AbnormalRes)(Artifact is HighArtifact) (1)
13. If (CovertCS is LowProportion) and (OvertCS is HighProportion) and (Gain is NormalGain) then (ResNormality is NormalRes)(Artifact is HighArtifact) (0.5)
14. If (CovertCS is LowProportion) and (OvertCS is LowProportion) and (Gain is LowGain) then (ResNormality is AbnormalRes)(Artifact is LowArtifact) (0.2)
15. If (CovertCS is HighProportion) and (OvertCS is LowProportion) and (Gain is NormalGain) then (ResNormality is NormalRes)(Artifact is HighArtifact) (0.5)

Figure 4: Rule list.



(a) CUS proportion to normality

(b) CUS proportion and gain to normality

(c) CUS proportion to artifact index

(d) CUS proportion and gain to artifact index

Figure 5: Membership function for each input and output. (a) Covert and overt CUS proportion. (b) Gain, (c) normality, (d) artifact.

second is Recommendation 2, with two inputs: position gain and appearance of covert CUS. The purpose of the design for the Recommendation 1 application is to reduce the confusion of variety in VOR gain. The second recommendation, Recommendation 2, will give the clinician a different output after combining CUS occurrence and position gain. Recommendation 2 requires one more input of covert CUS detection. The obtained normality, however, just somehow represents the effect of covert CUS on the diagnostic decision. Consequently, the new system was built to overcome the drawback

Table 1: System evaluation with the data of two controls and both sides.

Number	Covert CUS	Overt CUS	Gain	Normality	AI
Normal1_L	0.00	0.00	1.23	0.92	0.36
Normal1_R	0.19	0.05	1.10	0.96	0.10
Normal2_L	0.00	0.00	1.07	0.99	0.04
Normal2_R	0.18	0.05	1.14	0.94	0.16

Table 2: System evaluation with the data of 15 VN subjects.

Side number	Lesion					Intact				
	Covert CUS	Overt CUS	Gain	Normality	AI	Covert CUS	Overt CUS	Gain	Normality	AI
1	0.63	1.00	0.48	0.19	0.62	0.36	0.39	1.19	0.65	0.53
2	0.19	0.56	0.45	0.45	0.49	0.40	0.27	0.53	0.68	0.34
3	0.73	0.00	0.72	0.53	0.82	0.09	0.00	0.68	0.98	0.01
4	0.33	1.00	0.58	0.34	0.81	0.00	0.09	0.95	1.00	0.00
5	0.31	0.47	0.84	0.63	0.53	0.55	0.21	0.93	0.64	0.59
6	0.39	0.57	0.95	0.50	0.70	0.55	0.18	1.08	0.64	0.59
7	0.74	1.00	0.45	0.13	0.54	0.41	0.15	0.72	0.79	0.32
8	0.86	0.96	0.59	0.08	0.78	0.27	0.09	0.98	0.93	0.11
9	1.00	0.52	1.04	0.32	0.89	0.42	0.03	1.22	0.73	0.51
10	0.32	1.00	0.37	0.23	0.47	0.45	0.21	0.63	0.70	0.40
11	0.40	1.00	0.92	0.35	0.92	0.00	0.00	1.09	0.98	0.05
12	0.50	0.92	1.05	0.30	0.90	0.20	0.20	1.07	0.93	0.11
13	0.04	0.15	0.17	0.56	0.14	0.09	0.13	0.74	0.99	0.01
14	0.95	0.05	0.92	0.50	0.93	0.61	0.10	0.96	0.60	0.70
15	0.08	0.50	0.82	0.63	0.55	0.11	0.28	1.02	0.90	0.15
	Mean			0.38	0.67				0.81	0.29

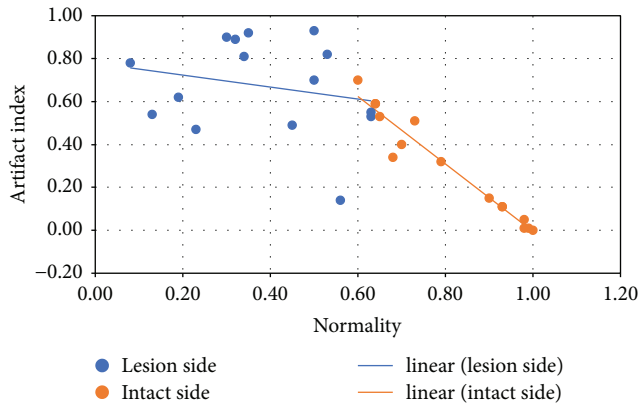


Figure 6: Artifact index versus normality in lesion side and intact side of 15 VN subjects.

of the previous system, using new inputs and new membership function, and produces one more factor of artifact index. The membership functions were built for three inputs: gain, proportion of covert catch-up saccades, and proportion of overt catch-up saccade as shown in Figure 5.

3.2. System Evaluation. Table 1 shows the evaluations from the dataset of two controls with both sides, and Table 2 shows the evaluation's result from the dataset of 15 VN subjects. In VN patient's population, in the lesion side, the normality is 0.38 ± 0.17, with the artifact index of 0.67 ± 0.21, for the intact side, the normality is 0.81 ± 0.29, with the artifact index of 0.29 ± 0.24; whereas for normal subjects, the normality is 0.95 ± 0.03, with the artifact index of 0.16 ± 0.34. The linear relation of artifact index and normality in VN population was plotted in Figure 6. This relationship is more random on the lesion side, whereas it is more linear on the intact side.

The artifact index appears to be very high in the lesion side of VN patients, indicating highly theoretical contradictions, which are low gain but without CUS, or normal gain with the appearance of CUS. Both intact sides and normal subject show high normality and low artifact index, even though the intact side has slightly lower normality and higher artifact index. The fifth and fifteenth subjects' lesion side reponses both have the highest normality of 0.63 among the lesion side's data, with relatively low artifact indices of 0.53 and 0.55, correspondingly. These results might indicate the lower severe condition compared to others, because of the high normality and low artifact index.

In addition, this program also gives out the optima point table, which could help the examiner review the maxima and minima point in the dataset with the number of optima points, their timing, and amplitude. The one with different value from the group could be the inference to gain calculation, hence should be removed.

4. Discussion

The software was built with a friendly graphical user interface (GUI), including plot, descriptive statistics for head and eye onset and zero crossing, gain, CUS timing and amplitude, and so on. In addition, it gave out an interface with the statistics of gain and CUS, plot of normal, covert and overt CUS-contained impulses, and normality and artifact. The whole GUI is mentioned in Figures 7–9. Normality represents the response's status by combining CUS proportion and gain using fuzzy inference system. The artifact index is an output from the fuzzy inference system, in which the impulse is contradicting with theory, that is, low gain but no CUS, or gain above 0.8 but still have CUS(s). The artifact index seems to be associated with the pattern of the whole dataset, which cannot be removed manually, and the clinician should

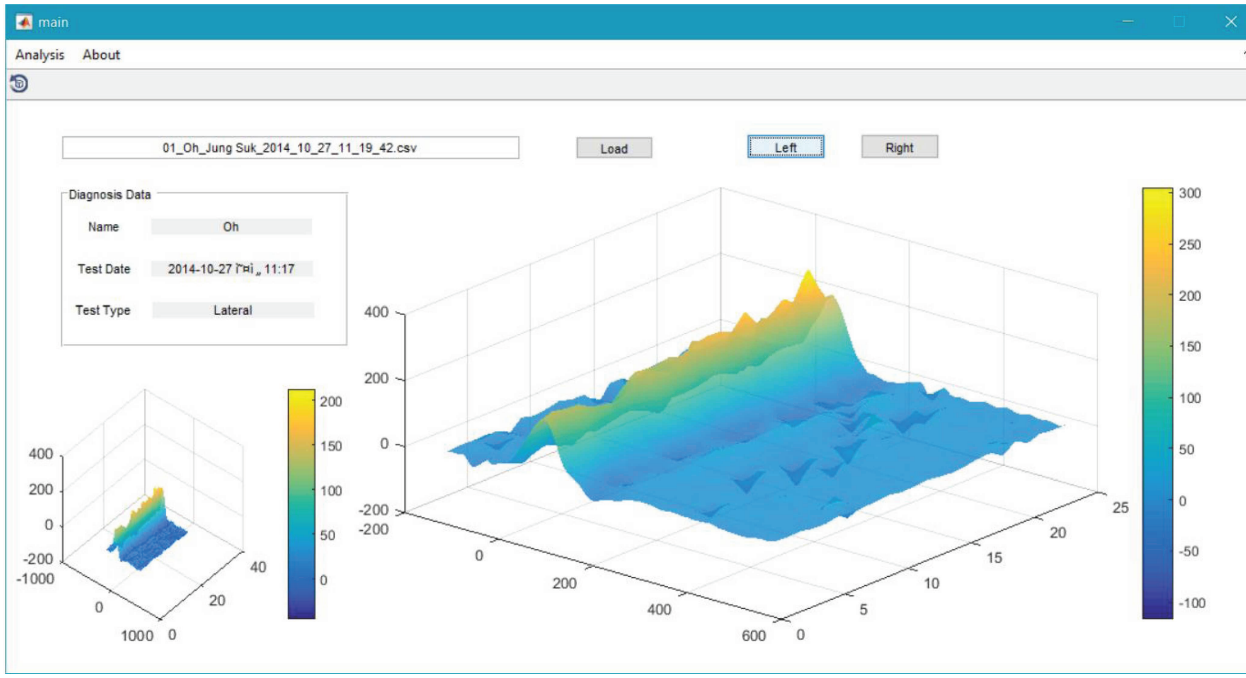


Figure 7: Main GUI of one side of one subject. The user can load the subject data on each side using "Load," "Left," and "Right" buttons. In this figure, the user can also find the metadata of the examination such as the subject's name, test date, and test type. In addition, the small plot is representing the head impulse, while the large one is representing the corresponding eye response.

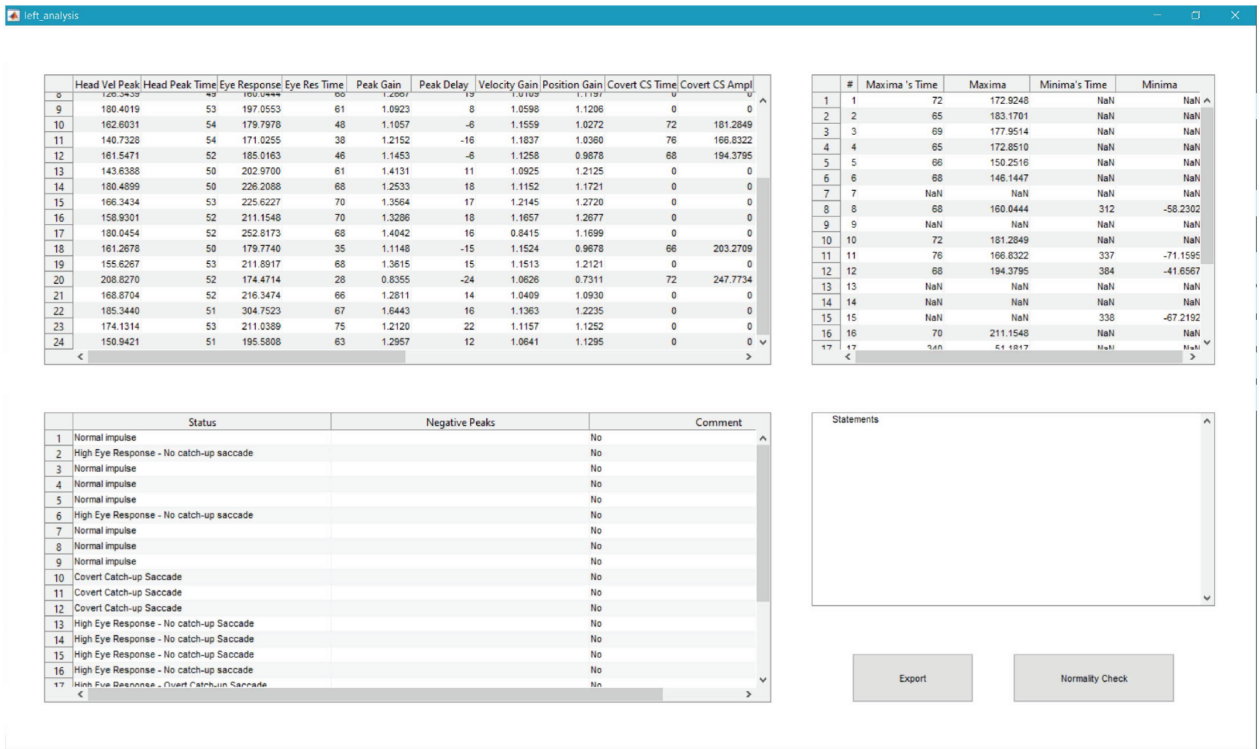


Figure 8: Statistical table of 14 parameters of one subject's left side. These parameters include head velocity peak's amplitude and time, eye response peak's value and time, peak gain (the ratio of eye response's peak to velocity impulse's peak), peak delay (delay of eye response's peak to head impulse's peaks), velocity VOR gain, position VOR gain, CCS amplitude and time, OCS amplitude and time, and rebound peak's amplitude and time. This figure also contains the optimal value table of all the significant minimum and maximum in the dataset. In addition, one descriptive table of status, which indicates that the response is normal, or contains catch-up saccade, is included. The user can note down statement and save for later use.

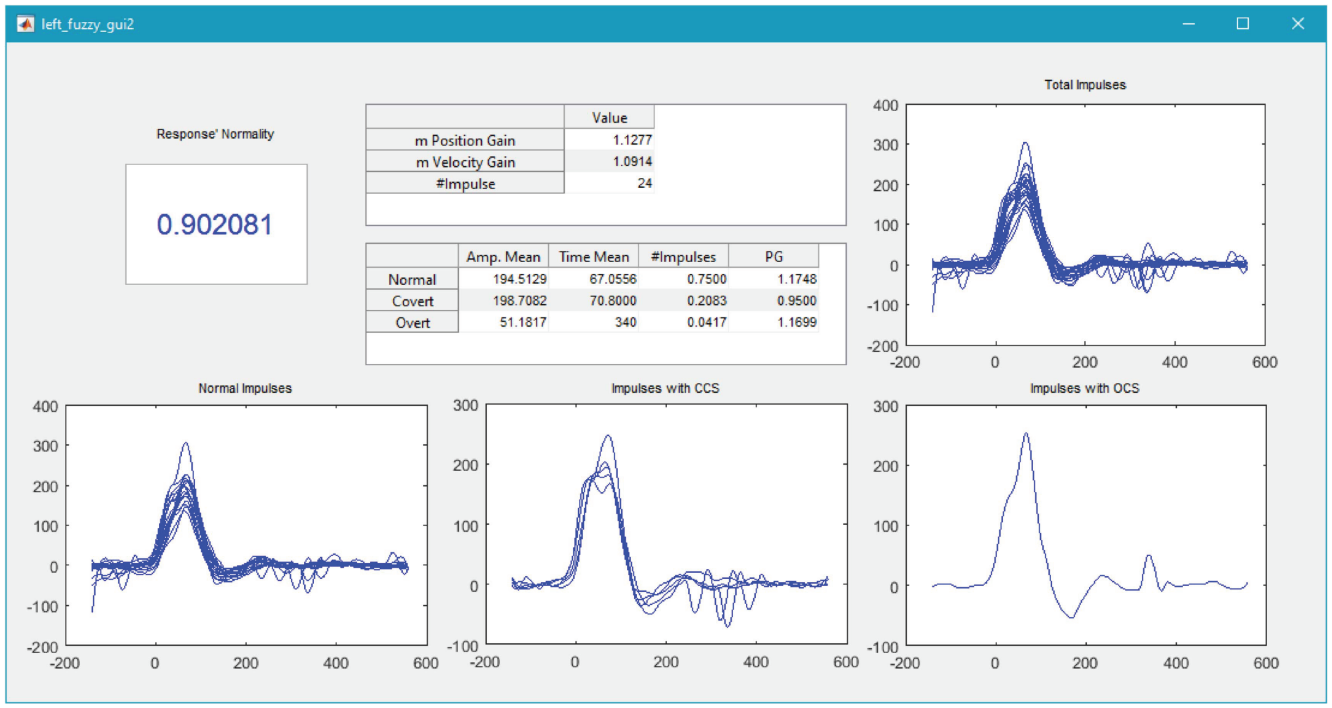


Figure 9: Normality and artifact index of the eye responses. This figure is evoked by clicking on the "Normality Check" button in the figure mentioned in Figure 8. In this figure, the user can find the normality, statistical analysis of gain, and catch-up saccades. This figure also includes four plots of total impulses, normal impulses only, impulses with CCS, and impulses with OCS.

reperform the test or adjust the normality of the dataset due to his/her own knowledge.

This is a robust system, which is the first one takes gain and CUS into account, to output not only the normality of the vHIT dataset, but also the artifacts. For future work, this system could be developed to work with any type of file output until it can get the head and eye movements with the right window. The current code works with ICS device only. We will provide the code as an open source. Furthermore, with sufficient data, this system can be highly developed with adaptive neural-network fuzzy inference system, for both normality and artifact. In addition, with clear understanding of the clustering of saccadic pattern as an input, this system can be developed further to apply in vestibular rehabilitation.

5. Conclusions

We successfully built the CDSS for vHIT VOR estimation using fuzzy inference system and evaluated the controls and VN dataset. This is a knowledge-based system that mimics the clinician's reasoning to output the normality and artifact within the vHIT dataset.

The vHIT is interfered by several factors coming from real practice such as examination environment and patient condition. In addition, the evaluation of vestibular function with vHIT is only reliable when the artifact is low. Therefore, in order to effectively examine the vHIT dataset, this CDSS system produces two outputs, the normality and the artifact index. So if the artifact index is low, the clinician can rely on the normality. If not, he/she can review the dataset and give her/his own opinion of the test. Therefore, the severe level is based on the normality and artifact. The recommended check protocol is firstly check for the artifact and the saccadic pattern, then evaluate and adjust the normality.

Conflicts of Interest

The authors declare that they have no conflict of interest.

Acknowledgments

This work was supported by the WCSL (World Class Smart Lab) program funded by Inha University.

References

[1] R. J. Leigh and S. Khanna, "Neuroscience of eye movements," *Advances in Clinical Neuroscience and Rehabilitation*, vol. 5, no. 6, 2006.

[2] I. S. Curthoys, H. G. MacDougall, L. A. McGarvie et al., "The video head impulse test (vHIT)," in *Balance Function Assessment and Management*, G. P. Jacobson and N. T. Shephard, Eds., pp. 391–430, Plural Publishing, San Diego, CA, USA, 2014.

[3] P. Cremer, G. Halmagyi, S. Aw et al., "Semicircular canal plane head impulses detect absent function of individual semicircular canals," *Brain*, vol. 121, no. 4, pp. 699–716, 1998.

[4] F. D. Roy and R. D. Tomlinson, "Characterization of the vestibulo-ocular reflex evoked by high-velocity movements," *Laryngoscope*, vol. 114, no. 7, pp. 1190–1193, 2004.

[5] K. P. Weber, S. T. Aw, M. J. Todd, L. A. McGarvie, I. S. Curthoys, and G. M. Halmagyi, "Head impulse test in unilateral vestibular loss: vestibulo-ocular reflex and catch-up saccades," *Neurology*, vol. 70, no. 6, pp. 454–463, 2008.

[6] H. G. MacDougall, K. P. Weber, L. A. McGarvie, G. M. Halmagyi, and I. S. Curthoys, "The video head impulse test: diagnostic accuracy in peripheral vestibulopathy," *Neurology*, vol. 73, no. 14, pp. 1134–1141, 2009.

[7] H. G. MacDougall, L. A. McGarvie, G. M. Halmagyi, I. S. Curthoys, and K. P. Weber, "The video head impulse test (vHIT) detects vertical semicircular canal dysfunction," *PLoS One*, vol. 8, no. 4, article e61488, 2013.

[8] A. Palla, D. Straumann, and H. Obzina, "Eye-position dependence of three-dimensional ocular rotation-axis orientation during head impulses in humans," *Experimental Brain Research*, vol. 129, no. 1, pp. 127–133, 1999.

[9] K. P. Weber, H. G. MacDougall, G. M. Halmagyi, and I. S. Curthoys, "Impulsive testing of semicircular-canal function using video-oculography," *Annals of the New York Academy of Sciences*, vol. 1164, no. 1, pp. 486–491, 2009.

[10] G. Mantokoudis, A. S. Saber Tehrani, J. C. Kattah et al., "Quantifying the vestibulo-ocular reflex with video-oculography: nature and frequency of artifacts," *Audiology & Neurotology*, vol. 20, no. 1, pp. 39–50, 2015.

[11] A. S. Hussein, W. M. Omar, X. Li, and M. Ati, "Accurate and reliable recommender system for chronic disease diagnosis," in *Global Health 2012: The First International Conference on Global Health Challenges*, pp. 113–118, Venice, Italy, 2012, IARIA.

[12] K. Kawamoto, C. A. Houlihan, E. A. Balas, and D. F. Lobach, "Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success," *BMJ*, vol. 330, no. 7494, p. 765, 2005.

[13] T. A. D. Nguyen, D. Y. Kim, S. M. Lee, K. S. Kim, S. R. Lee, and J. W. Kwon, "Recommendation application for video head impulse test based on fuzzy logic control," *Journal of Central South University*, vol. 23, no. 5, pp. 1208–1214, 2016.

[14] A. Blödow, S. Pannasch, and L. E. Walther, "Detection of isolated covert saccades with the video head impulse test in peripheral vestibular disorders," *Auris Nasus Larynx*, vol. 40, no. 4, pp. 348–351, 2013.

[15] S. J. Herdman, "Vestibular rehabilitation," in *Contemporary Perspectives in Rehabilitation*, p. 18, F.A. Davis, Philadelphia, PA, USA, 2007.

[16] J. Rey-Martinez, A. Batuecas-Caletrio, E. Matiño, and N. Perez Fernandez, "HITCal: a software tool for analysis of video head impulse test responses," *Acta Oto-Laryngologica*, vol. 135, no. 9, pp. 886–894, 2015.

[17] R. Miles, *Vorteq™ vHIT. White Paper*, Micromedical Technologies, Chatham, IL, USA, 2013.

[18] W. H. Press, S. Teukolsky, W. T. Vetterling, and B. P. Flannery, "Savitzky-Golay smoothing filters," in *Numerical Recipes in C: the Art of Scientific Computing*, pp. 650–655, Cambridge University Press, New York, NY, USA, 1988.

[19] J. M. Mendel, "Fuzzy logic systems for engineering: a tutorial," *Proceedings of the IEEE*, vol. 83, no. 3, pp. 345–377, 1995.

[20] F. Dernoncourt, *Fuzzy Logic: Introducing Human Reasoning within Decision-Making Systems?*, Conservatoire National des Arts et Métiers, Paris, France, 2011.

[21] J. T. Ross, *Fuzzy Logic with Engineering Applications*, John Wiley & Sons, Ltd, Chichester, UK, 2010.

[22] J. W. Hines, *MATLAB Supplement to Fuzzy and Neural Approaches in Engineering*, John Wiley & Sons, Inc, New York, NY, USA, 1997.

Hindawi
Journal of Sensors
Volume 2018, Article ID 4628245, 10 pages
https://doi.org/10.1155/2018/4628245

Research Article

Mcredit2: Enhanced High-Performance Xen Scheduler via Dynamic Weight Allocation

Minsu Kang and Sangjun Lee

Department of Computer, Soongsil University, Seoul 156-743, Republic of Korea

Correspondence should be addressed to Sangjun Lee; sangjun@ssu.ac.kr

Received 31 October 2017; Accepted 22 January 2018; Published 26 March 2018

Academic Editor: Mucheol Kim

Generally, operating only a single host on a single server results in hardware underutilization. Thus, hypervisors (e.g., Xen) have been developed to allow several hosts to operate on a single server. The Xen hypervisor provides processor schedulers (e.g., Credit and Credit2 schedulers) to assign processors to each host. The Credit2 scheduler provides work assurance to the domain relative to latency and it evenly assigns processors to each domain. In addition, the Credit2 scheduler can assign a weight value to each host. A greater host weight value allows processors to be assigned to a host for longer periods. However, the Credit2 scheduler shows poorer performance than the basic Credit scheduler, which utilizes idle processors. In this paper, we propose the Mcredit2 scheduler, which improves the Credit2 scheduler. The Credit2 scheduler takes no action when the load on a specific domain causes increased processor usage. The proposed Mcredit2 scheduler allows a domain to quickly process loads by temporarily assigning a greater weight value to a host with high processor usage. In addition, we introduce a processor monitoring tool that visualizes the processor usage.

1. Introduction

In recent years, with the significant advances in computer and communication technologies, which are new big leaps in the digital world, a term called "Internet of Things" (IoT) made its appearance [1]. Only in 2011 did the number of interconnected devices on the planet overtake the actual number of people, and it is expected to reach 24 billion devices by 2020 [2]. In such an environment, IoT servers are required to efficiently process large-scale data collected from those devices. Generally, using only a single host on a high-availability server can result in underutilized hardware. Table 1 shows the hardware performance of an IBM server [3]. To minimize hardware underutilization, hardware virtualization techniques [4, 5] have been studied. Server virtualization [6, 7] techniques allow several hosts to operate on a single server. The use of hypervisors [8, 9] to support server virtualization and reduce hardware underutilization has also been studied. Currently, hypervisors, such as KVM [10] and Xen [11], have gained popularity.

With Xen, the host is called the domain, and Xen requires a scheduler to assign processors to each domain. The scheduler allocates a physical processor to the virtual processors of each domain to perform operations.

Borrowed Virtual Time, Atropos, and Round Robin schedulers [12] were used in early versions of Xen. Currently, Xen 4.5.0 supports the Simple Earliest Deadline First (SEDF) [13], Credit [14], Credit2 [15], Real-time Deferrable Server [16], and Avionics Application Standard Software Interface [17] schedulers. The Credit scheduler is commonly used among these schedulers.

The Xen hypervisor adopts the Credit2 scheduler beginning with Version 4.6 [17]. The Credit2 scheduler can assign a weight value to each domain. The higher the weight value, the longer the time for which processors can be assigned. In this paper, we introduce the Mcredit2 scheduler as an improvement to the Credit2 scheduler. The Mcredit2 scheduler temporarily allocates a greater weight value to a domain having higher processor utilization. This allows domains that suddenly require many processors (e.g., web servers) to

Table 1: System x3950 X6 specifications [18].

Processor	Up to eight Intel Xeon E7-8800 v2 processors, each in a compute book. Each processor has 15 cores (up to 2.8 GHz), 12 cores (up to 2.6 GHz), 10 cores (up to 2.2 GHz), eight cores (up to 2.0 GHz), or six cores (up to 3.4 GHz). There are three QPI links (up to 8.0 GTps each). Up to 1600 MHz memory speed. Up to 37.5 MB L3 cache. Intel C602J chipset.
Memory	Up to 192 DIMM sockets (24 DIMMs per processor, installed in the compute book). RDIMMs and load-reduced DIMMs are supported, but memory types cannot be intermixed. Memory speed up to 1600 MHz.
Memory maximums	With RDIMMs: up to 3 TB with 192 × 16 GB RDIMMs and eight processors. With LRDIMMs: up to 12 TB with 192 × 64 GB LRDIMMs and eight processors.
Memory protection	ECC, Chipkill, RBS, memory mirroring, and memory rank sparing.

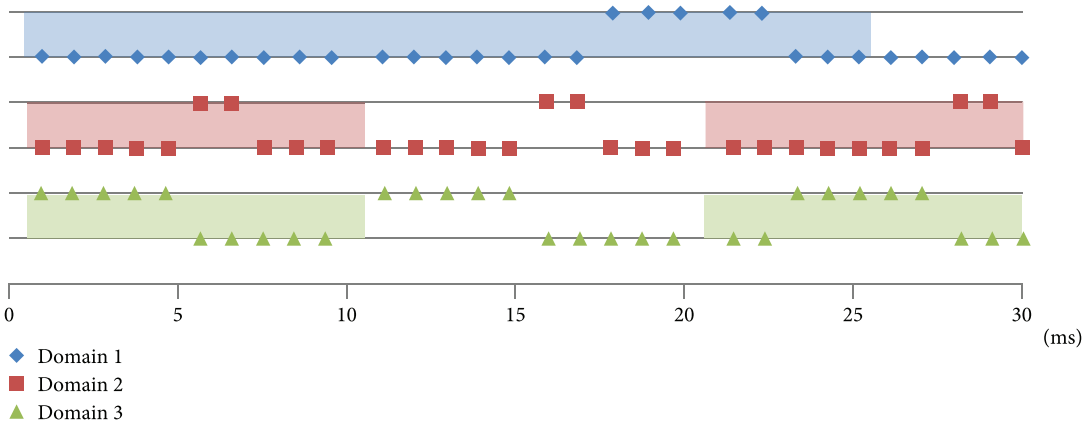


Figure 1: SEDF scheduler CPU allocation.

operate more flexibly. In this paper, we extend our previous work [19] and present more experimental results and intensive analysis that were not reported in the previous work.

The server administrator can allocate domain resources by checking whether the current CPU usage per domain is high. It can also check the average CPU usage in the domain for a certain unit time. The Xen hypervisor provides a monitoring tool such as xentop [20] that shows information about the Xen system in real time. However, with xentop, only the usage time of the current CPU is indicated with text, creating difficulties for a server administrator. In this paper, we propose a virtual processor monitoring tool called Mcredit2 that visually displays the CPU usage for each domain and shows the CPU usage time, usage percentage, and average usage time.

The remainder of this paper is organized as follows. In Section 2, the basic Xen hypervisor scheduler is described. In Section 3, the proposed scheduler and monitoring tool implementation are described. In Section 4, scheduler performance is evaluated based on the implementation results. Finally, the paper is concluded in Section 5.

2. Backgrounds

Here, the SEDF, Credit, and Credit2 schedulers are described to facilitate a comparison with the Mcredit2 scheduler.

2.1. SEDF Scheduler. The SEDF scheduler employs a real-time scheduling method [21], that is, the scheduler assigns higher priority to the process having the earliest deadline.

Each domain has a period parameter and a slice parameter. Each domain is executed for nanoseconds of the slice parameter value and every nanosecond of the period parameter value. These parameters are used to determine the domain having the earliest deadline.

For example, the assumed domains 1, 2, and 3 have period and slice parameters of 25 and 5, 10 and 2, and 10 and 5, respectively. Domains 2 and 3 have the shortest deadlines, that is, a period parameter of 10. Here, domain 2 will be executed within 5 ms and domain 3 will wait for 8 ms. Therefore, domain 3 is scheduled first. Domain 2's deadline becomes shorter than that of domain 3 after scheduling. After domains 2 and 3 are scheduled for 20 ms, domain 1 is scheduled for 5 ms. Figure 1 shows a diagram of this scheduling process. In Figure 1, the horizontal axis shows the elapsed time and the background color changes depending on the period parameter. The increase of the value reflects scheduling.

2.2. Credit Scheduler. Currently, the Credit scheduler [22] is the basic Xen scheduler. The Credit scheduler allocates symbolic money called credit to domains. Each domain pays credits to use processors. The Credit scheduler ensures that all work is processed. Even if all credits assigned to a domain are used, additional processors can be utilized if there are idle processors.

With the Credit scheduler, each domain has a weight value parameter and a capacity parameter. The weight value determines when a domain can use processors. Credits are assigned to a domain in proportion to the weight value. For

example, if domain 1 has a weight value of 512 and domain 2 has a weight value of 256, twice the number of credits are assigned to domain 1. As a result, domain 1 processors can be used for twice as long as domain 2 processors. The default weight value is 256, the minimum is 1, and the maximum is 65,535. The capacity parameter determines the upper limit of the processors to be used in the domain. A capacity value of zero means that there is no upper limit. The upper limit of one processor increases each time the capacity is increased by 100. The processor utilization cannot be over capacity, even with idle processors.

The basic Credit scheduler algorithm is as follows. Credits are assigned to all domains every 30 ms relative to the weight value. The virtual processors using processors pay credits every 10 ms. With the Credit scheduler, virtual domain processors are assigned BOOST, UNDER, OVER, and IDLE states. Processors in the BOOST state are given the highest priority, followed by UNDER, OVER, and IDLE states. The processor enters the BOOST state when an idle virtual processor attempts to process a task. A virtual processor enters the UNDER state when using a physical processor. The processor enters the UNDER state when the virtual processors have remaining credits. The processor enters the OVER state when the virtual processors do not have any credits. When a virtual processor is not in use, it enters the IDLE state. The state of the virtual processors determines the priority queue assigned to each according to priority.

The Credit scheduler has limitations. For example, it saves remaining credits and uses them for subsequent processing tasks. If the credits of a given domain are greater than a certain value, that domain becomes inactive and the unused credits are discarded. If the domain is executing multimedia, processors remain idle much of the time, because they are only required when decoding multimedia. The domain's virtual processors enter the BOOST state and are positioned at the top of the priority queue when decoding. After decoding, they enter the UNDER state and compete with other virtual processors. These processors rank far below other processors in priority and are decoded erratically, because the domain state is changed to inactive and saved credits are discarded. As a result, multimedia playback may be terminated.

2.3. Credit2 Scheduler. The Credit2 scheduler [11] is intended to address the weaknesses of the Credit scheduler. Like the Credit scheduler, the Credit2 scheduler provides credits to domains to use processors, and each domain must pay credits to use the processors. The Credit2 scheduling algorithm differs from that of the Credit scheduler.

The Credit2 scheduler algorithm is described as follows. The same credits are assigned to all domains and virtual processors pay credits when using a processor. In this case, the amount of credits paid by virtual processors varies with the weight value of the domain, even if processors are used for the same amount of time. A greater domain weight value results in lower credit costs. Concurrently, the credits to be paid increase in proportion to the greatest domain weight value. For the domain with the greatest weight value, processors can be used for 10 ms and all credits are used. Regarding the priority assigned to each virtual processor, the more credits remaining, the higher the priority becomes. When a specific virtual processor uses all its credits, the credit value of all virtual processors will be initialized.

Another difference between the Credit2 and Credit schedulers is how virtual processors are prioritized in the queue. The Credit scheduler has one priority queue per processor, and one processor schedules several virtual processors. With the Credit2 scheduler, there is one priority queue for each set of processors.

2.4. Xentop. Xen provides a monitoring tool called xentop [20] that operates as shown in Figure 2.

It is difficult to check the amount of usage of virtual processors in real time because xentop reflects only their execution time. Improved monitoring tools are required for an administrator to check the usage amount of the virtual processors for each domain.

3. Proposed Techniques

In this section, we describe the Mcredit2 scheduler and the implementation of the virtual processor monitoring tools.

3.1. Mcredit2 Scheduler. The Mcredit2 scheduler dynamically changes the weight value depending on the given situation. In contrast, the weight value is only changed by an administrator with the Credit2 scheduler. With the Mcredit2 scheduler, for an insufficient number of credits, the initialization of credits for all processors is executed when virtual processors attempt to use physical processors. At this time, a variable is assigned to the domain that has virtual processors with insufficient credits to increase the variable. At fixed intervals, the weight value of the domain with the greatest variable is increased. This weight value increases during the BOOST state. If the number of processors for a specific domain increases, the virtual processors in that domain will not have sufficient credits and the initialization of credits increases. Therefore, the domain can perform the given job with an increased weight value.

Figure 3 shows the structures used in the Mcredit2 scheduler, such as the mcsched_runqueue_data, mcsched_dom, and mcsched_vcpu structures. Here, runq in the mcsched_runqueue_data structure is a list of viable virtual machines, and svc is a list of all virtual processors in that runqueue. The max_weight saves the greatest weight value in vcpu included in the runqueue. The vcpu in the mcsched_dom structure is a list of vcpu of the domain. In addition, weight is the weight value of the domain and the vcpu of the domain. The nr_vcpus stores the number of vcpu of that domain, and reset_count increases the value each time the vcpu of the domain initializes with credits. The sdom in the mcsched_vcpu structure refers to a domain that contains itself. Here, weight stores the value required for the algorithm used to reduce the number of credits. Credit stores the credit value to be paid when vcpu uses cpu.

Figure 4 shows how to enter the BOOST state after executing the reset_credit() function to initialize credits for all vcpus because a specific vcpu has insufficient credits. Here, sdom is the mcsched_dom structure having vcpu to call the

```
xentop - 15:14:24     Xen 4.5.0
5 domains: 1 running, 4 blocked, 0 paused, 0 crashed,0 dying, 0 shutdown
Mem: 20664348k total, 5519256k used, 15145092k free            CPUs: 8 @ 3400MHz
      NAME   STATE    CPU (sec)  CPU (%)     MEM (k)  MEM (%)   MAXMEM (k)  MAXMEM (%)  VCPUS
 NETS  NETTX (k)  NETRX (k)  VBDS     VBD OO     VBD RD     VBD WR                VBD WSECT  SS
ID
  Domain-0 -----r           223     1.1      1048576     5.1    no limit         n/a      8
    0         0         0     0         0          0          0           0           0
0
  Domain_1 --b---             9     0.0      1048576     5.1      1049600         5.1      4
    1       521       185     1         0       3179        179       185710        3152
0
  Domain_2 --b---            95    31.7      1048576     5.1      1049600         5.1      2
    1        40         3     1         0       3339        208       185494        5144
0
  Domain_3 --b---           112    37.1      1048576     5.1      1049600         5.1      2
    1        40         3     1         0       3287        208       193718        5224
0
  Domain_4 --b---            48    15.9      1048576     5.1      1049600         5.1      2
    1        35         3     1         0       3282        214       185278        5136
0

Delay  Networks  vBds  Tmem  VCPUs  Repeat  header  Sort  order  Quit
```

Figure 2: Xentop execution screen.

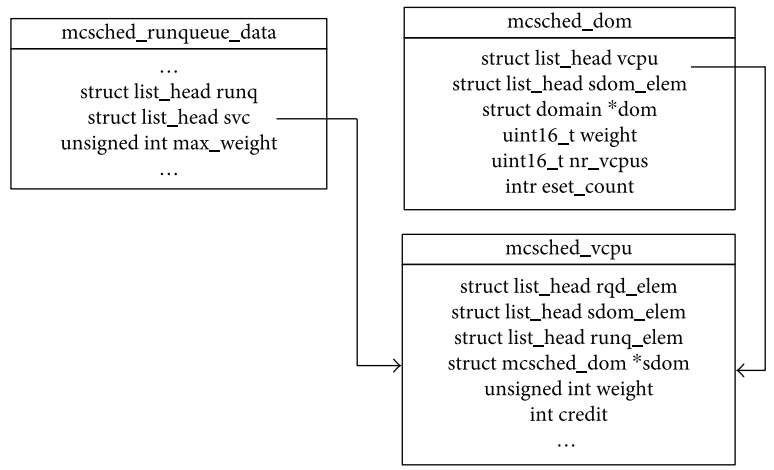


Figure 3: mcsched_runqueue_data, mcsched_dom, and mcsched_vcpu structures.

reset_credit() function owing to insufficient credits. The time at which the reset_credit() function is called and the domain of BOOST state is changed by checking tick variables. A domain in the BOOST state is changed after sufficient time has passed. Otherwise, the value of the reset_count variable of sdom is increased by one. If this reset_count value happens to be the greatest, the domain to be changed to the BOOST state next is reserved as sdom.

Figure 5 shows a flowchart of the process to change the BOOST state domain in the routine for initializing credits. Here, old_d variable is the mcsched_dom structure holding the current BOOST state domain. The d variable is the mcsched_dom structure of the domain to enter the BOOST state. The rate variable saves the increase rate of the current weight value. If the current BOOST state domain and the domain to be changed to BOOST state are the same, the rate increases two times. Otherwise, the weight value is initialized by dividing the current BOOST state domain by the rate. Then, the current BOOST state domain is updated by changing old_d to d after initializing the rate to 2. Finally, multiplying the original weight value of the current BOOST state domain by the rate value allows the weight value to be dynamically changed.

Algorithm 1 shows the process for assigning the BOOST state to a domain and its initialization.

3.2. Virtual Processor Monitoring Tools. The xenstat_domain and xenstat_vcpu structures of Figure 6 are declared in

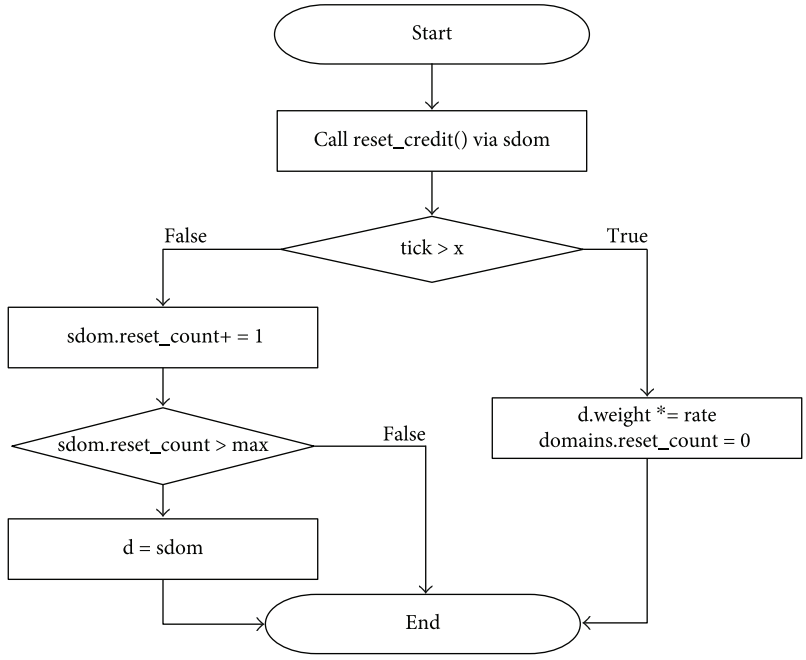


Figure 4: BOOST procedure.

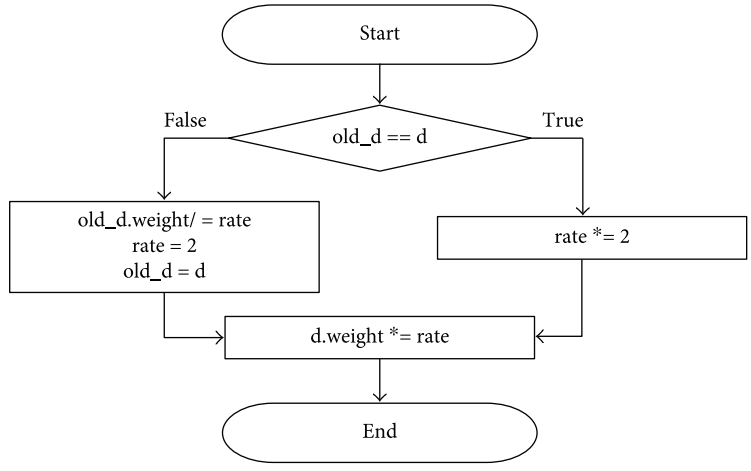


Figure 5: BOOST initialization process.

xenstat_priv.h. cpu_ns in the xenstat_domain structure is a variable that stores the total use time of a domain's processor in nanoseconds. num_vcpus is a variable that stores the number of virtual processors assigned to the domain, and vcpus is a variable with an xenstat_vcpu structure address created using the allocated virtual processors. Here, ns in the xenstat_vcpu structure is a variable that stores the use time of a virtual processor in nanoseconds.

The improved monitoring tools provide the usage of the virtual processors assigned to each domain using the variables described above as percentages. The implementation of the monitoring tools is described as follows.

Essentially, the program saves the xenstat_node structures used previously to calculate the virtual CPU utilization. The time is obtained in microseconds using the previous and the current output time. The use percentage of the virtual processors is obtained using the ns variable of the xenstat_vcpu structure. This process is similar to that shown in Figure 7.

The calculation of virtual CPU utilization (%) is given in Algorithm 2.

Figure 8 shows the complete operation of the proposed virtual processor monitoring tools. The virtual processor utilization for all domains is evaluated using these tools. These tools visually represent the utilization of a virtual CPU for each domain in bar form depending on the utilization specified by each hashtag. The tools also show the average utilization of virtual CPUs.

```
1:  begin
2:    Input: mcsched_dom* snext that calls the reset_credit() function
3:    Output: execution time of the snext->vcpu and snext
4:    // timed wait and enter BOOST state.
5:    if tick > n then
6:      // previous BOOST state: mcsched_dom prev_boost_sdom.
7:      // initialize a weight value
8:      for all vcpus ∈ prev_boost_sdom do
9:        vcpus→weight ← vcpus→weight/rate;
10:     end
11:     prev_boost_sdom→weight ←
                                  prev_boost_sdom→weight/rate;
12:     // future BOOST state: mcsched_dom boost_sdom.
13:     // Increment of the magnification of the weight value.
14:     if prev_boost_sdom = boost_sdom then
15:       rate ← rate*2;
16:       // Initialization of the magnification of the weight value
17:     else
18:       prev_boost_sdom ← boost_sdom;
19:       rate ← 2;
20:     end
21:     // Initialization of the reset_count variables for all domains
22:     for all vcpus ∈ rqd do
23:       svc→sdom→reset_count ← 0;
24:     end
25:     // Increase the weight value for the BOOST state domain.
26:     for all vcpus ∈ boost_sdom do
27:       vcpus→weight ← vcpus→weight*rate;
28:     end
29:     boost_sdom→weight ← boost_sdom→weight/rate;
30:   else
31:     // mcsched_vcpu snext that calls the reset_credit() function.
32:     // Increase the reset_count variable of mcsched_dom using snext
33:     snext→sdom→reset_count ←
                                  snext→sdom→reset_count+1;
34:     // update a mcsched_dom having the highest reset_count value
35:     if snext→sdom→reset_count > max_reset_count then
36:       max_reset_count ← snext_sdom→reset_count;
37:       boost_sdom ← snext→sdom;
38:     end
39:   end
40: end
```

Algorithm 1: BOOST state transition.

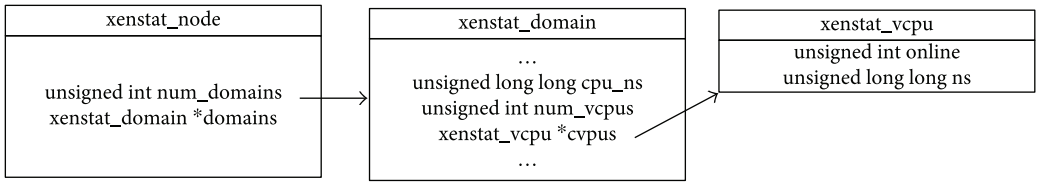


Figure 6: xenstat_node, xenstat_domain, and xenstat_vcpu structures.

4. Performance Evaluation

Here, the performance of the Mcredit2 scheduler is analyzed and compared to extant schedulers.

4.1. Experimental Environments. The experimental environment is as follows: Intel i7-2600 processor (four processors and eight threads) [23], 20 GB RAM, and a 1 TB hard disk. As a result, processor utilization for all domains can be up to 800%. Xen hypervisor version 4.5.0 is used. Ubuntu 14.04 LTS [24] is used as the domain operating system. A total of nine domains are generated. Four virtual processors, 1 GB RAM, and a 40 GB hard disk are assigned to each domain.

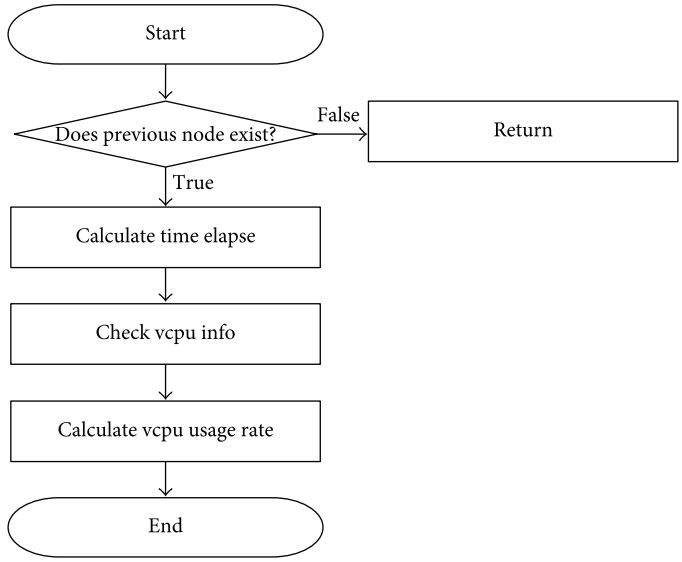


Figure 7: Calculation of virtual CPU utilization (%).

```
1:  begin
2:    Input: xenstat domain d
3:    Output: percentage of the virtual CPU utilization pct
4:    old_domain ← NULL;
5:    us_elapsed ← NULL;
6:    // check if previous node is exist
7:    if prev_node = NULL then
8:      return 0.0;
9:    else if vcpu = NULL then
10:     return 0.0;
11:   else if old_vcpu = NULL then
12:     return 0.0;
13:   end
14:   old_domain ←
      xenstat_node_domain(prev_node, xenstat_domain_id(domain));
15:   if old_domain = NULL then
16:     return 0.0;
17:   end
18:   // calculation of the elapsed time
19:   us_elapsed ←
            (curtime.tv_sec-oldtime.tv_sec)*1000000.0+
            (curtime.tv_usec-oldtime.tv_usec);
20:   // calculate the utilization percentage for the virtual CPU
21:   pct ←
            ((xenstat_vcpu_ns(vcpu)-xenstat_vcpu_ns(old_vcpu))/10.0)
            /us_elapsed;
22:   return pct;
23: end
```

Algorithm 2: Calculation of virtual CPU utilization (%).

The sysbench [25] benchmark tool is used to analyze the scheduler. The sysbench tool measures operating system performance using multithreads. In this experiment, the sysbench cpu mode is used as the test mode. In cpu mode, a request to calculate a decimal value is assigned to the generated thread. Four threads are generated for the sysbench tool. The maximum decimal value is 10,000 and the request to calculate the decimal value occurs 100,000 times. The average results of performing this process 10 times are obtained.

The environment variables for each scheduler are as follows. For the SEDF scheduler, the period and slice parameters of each domain are set to 100 and 10, respectively. For

```
       NAME   STATE    CPU (sec)  CPU (%)  AVR (%)
  Domain_2 --b---             17    30.9     32.7
VCPUS   #0,   15.2%  : [#######                    ]
VCPUS   #1,   15.8%  : [#######                    ]
       NAME   STATE    CPU (sec)  CPU (%)  AVR (%)
  Domain_3 --b---             16    36.9     38.6
VCPUS   #0,   20.8%  : [##########                 ]
VCPUS   #1,   16.1%  : [########                   ]
       NAME   STATE    CPU (sec)  CPU (%)  AVR (%)
  Domain_4 --b---              9    16.6     16.4
VCPUS   #0,    0.0%  : [                           ]
VCPUS   #1,   16.6%  : [########                   ]
```

Figure 8: Monitoring tool for virtual CPUs.

```
1:  begin
2:    Input: the number of threads: n
3:    Output: error number: err_num
4:    // infinite loop
5:    loop
6:      // impose load
7:      for var i < 10000000 do
8:        i ← i+1;
9:      end
10:     // stop to impose
11:   usleep(200000);
12:   end
13: end
```

Algorithm 3: Load test.

the Credit scheduler, the weight value and capacity of each domain are set to 256 and 0, respectively. For the Credit2 scheduler, the weight value of each domain is set to 256. For the proposed Mcredit2 scheduler, the weight value of each domain is set to 256.

The pseudocode of the program used to impose a transient load on each domain is shown in Algorithm 3.

In these experiments, the pseudocode threads are generated in the domain not to execute the sysbench benchmark tool. The domain that executes the sysbench benchmark is called Domain_1, and the other domains are named Domain_2 to Domain_9. In the first experiment, four threads are generated with sufficient idle processors, and in the second experiment, eight threads are generated with insufficient idle processors.

4.2. Experimental Results. Figures 9 and 10 show the total time required by the sysbench benchmark tool to measure performance and the idle processors generated during the process. Figure 9 shows the performance results with sufficient idle processors, and Figure 10 shows the results with insufficient idle processors.

The SEDF scheduler demonstrates the worst performance. It cannot give priority to Domain_1, which requires many virtual processors because the deadlines of all domains are the same.

For the Credit scheduler, even when a specific domain runs out of credits, if it has idle processors, the processors can be used without redistributing credits. This leads to increased virtual processor utilization of domains other than Domain_1. In addition, the Credit scheduler cannot use processors until credits are recharged when all credits are used. As a result, additional processors are not assigned to Domain_1 if there are no idle processors, and the performance results for Domain_1 are not as good as those of Credit2 and Mcredit2.

For the Credit2 scheduler, fewer processors are assigned to Domain_1 even if idle processors are sufficient. As a result, it shows poorer performance than the Credit scheduler relative to utilizing more idle processors.

The Mcredit2 scheduler can use more processors with the same number of credits because it dynamically increases the weight value in domains with significant load. If there are idle processors, processors can be continuously supplied due to the high weight value of Domain_1. Due to the greater weight value of Domain_1, it gains higher processor priority than the other domains and its performance is increased.

5. Conclusion

Hypervisors have been developed to take advantage of the increased performance of server hardware, and scheduling algorithms have also been improved, so that hypervisors can efficiently assign processors to virtual machines. As described in this paper, the Credit2 scheduler algorithm of the Xen hypervisor has been improved. In the proposed method, a domain that requires many processors under the Credit2 scheduler is identified, and the Mcredit2 scheduler gives higher priority to this domain by increasing its weight value.

The performance of the SEDF, Credit, and Credit2 schedulers were compared to that of the proposed Mcredit2 scheduler. The performance was measured using the sysbench benchmark tool. The SEDF scheduler shows significantly worse performance than the other schedulers, and the Credit scheduler uses idle processors. Therefore, higher virtual CPU utilization of domains can be accomplished with lower domain workload. With the Credit2 scheduler, when a load occurs in a specific domain, more virtual processors are assigned to that domain than with the Credit scheduler. With the proposed Mcredit2 scheduler, a domain where load occurs is assigned more processors. The experimental results demonstrate that the Mcredit2 scheduler assigns virtual

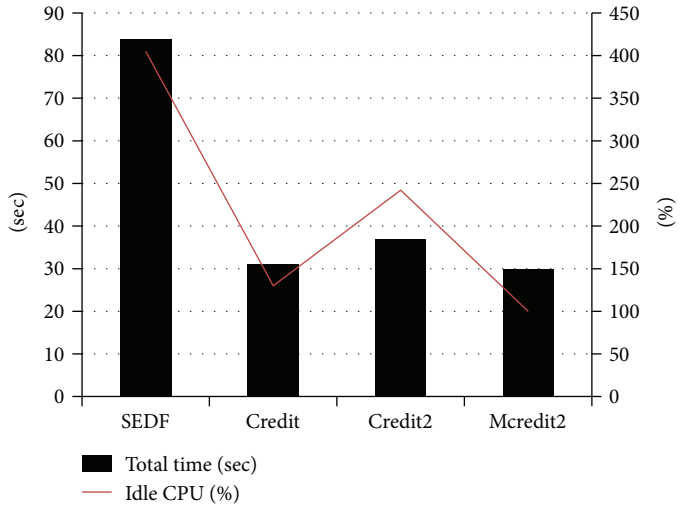


Figure 9: Results with sufficient idle CPUs.

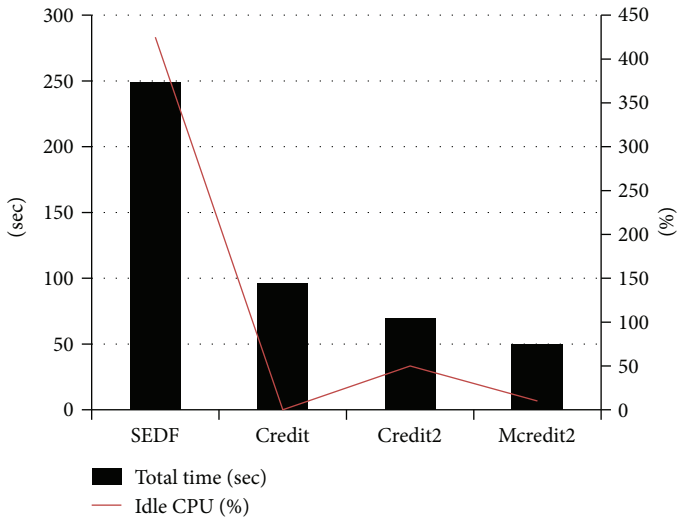


Figure 10: Results with insufficient idle CPUs.

processors to loaded domains more flexibly than the Credit2 scheduler. The experimental results also show that the Mcredit2 scheduler assigns processors more flexibly than the Credit2 scheduler when processor requirements for a specific domain significantly increase.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Korea Government (MSIP) (2015R1D1A1A01057680).

References

[1] A. P. Plageras, K. E. Psannis, C. Stergiou, H. Wang, and B. B. Gupta, "Efficient IoT-based sensor BIG Data collection–processing and analysis in smart buildings," *Future Generation Computer Systems*, vol. 82, pp. 349–357, 2018.

[2] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): a vision, architectural elements, and future directions," *Future Generation Computer Systems*, vol. 29, no. 7, pp. 1645–1660, 2013.

[3] November 2017, https://www.ibm.com/us-en/.

[4] M. F. Mergen, V. Uhlig, O. Krieger, and J. Xenidis, "Virtualization for high-performance computing," *ACM SIGOPS Operating Systems Review*, vol. 40, no. 2, pp. 8–11, 2006.

[5] L. Van Doorn, "Hardware virtualization trends," in *Proceedings of the 2nd international conference on Virtual execution environments '06*, vol. 14no. 16, ACM/Usenix, p. 45, New York, NY, USA, 2006.

[6] M. Steinder, I. Whalley, D. Carrera, I. Gaweda, and D. Chess, "Server virtualization in autonomic management of heterogeneous workloads," in *2007 10th IFIP/IEEE International Symposium on Integrated Network Management*, pp. 139–148, Munich, Germany, 2007.

[7] J. Daniels, "Server virtualization architecture and implementation," *Crossroads*, vol. 16, no. 1, pp. 8–12, 2009.

[8] R. P. Goldberg, "Survey of virtual machine research," *Computer*, vol. 7, no. 6, pp. 34–45, 1974.

[9] M. Rosenblum and M. Garfinkel, "Virtual machine monitors: current technology and future trends," *Computer*, vol. 38, no. 5, pp. 39–47, 2005.

[10] A. Kivity, Y. Kamay, D. Laor, U. Lublin, and A. Liguori, "kvm: the Linux virtual machine monitor," *Proceedings of the Linux Symposium*, vol. 1, 2007.

[11] P. Barham, B. Dragovic, K. Fraser et al., "Xen and the art of virtualization," *ACM SIGOPS Operating Systems Review*, vol. 37, no. 5, pp. 164–177, 2003.

[12] November 2017, https://wiki.xenproject.org/wiki/Xen_Project.

[13] D. Chisnall, *The Definitive Guide to the Xen Hypervisor*, Pearson Education, Upper Saddle River, NJ, USA, 2008.

[14] November 2017, https://wiki.xenproject.org/wiki/Credit_Scheduler.

[15] November 2017, https://wiki.xen.org/wiki/Credit2_Scheduler_Development.

[16] November 2017, https://wiki.xenproject.org/wiki/RTDS-Based-Scheduler.

[17] November 2017, https://wiki.xenproject.org/wiki/ARINC653_Scheduler.

[18] D. Watts and I. Krutov, *System x3950 X6 (3837)*, lenovopress, China, 2015.

[19] M. Kang, *Mcredit2: Enhanced Xen Scheduler for High Performance via Dynamic Weight Allocation [M.S. thesis]*, Soongsil University, Republic of Korea, 2016.

[20] November 2017, https://wiki.xenproject.org/wiki/Xentop.

[21] G. Lipari, *Earliest Deadline First*, Scuola Superiore Sant'Anna, Pisa-Italy, 2005.

[22] L. Cherkasova, D. Gupta, and A. Vahdat, "Comparison of the three CPU schedulers in Xen," *SIGMETRICS Performance Evaluation Review*, vol. 35, no. 2, pp. 42–51, 2007.

[23] December, http://ark.intel.com/products/52213/Intel-Core-i7-2600-Processor-8M–Cache-up-to-3_80-GHz.

[24] November 2017, https://wiki.ubuntu.com/TrustyTahr/ReleaseNotes.

[25] A. Kopytov, *SysBench Manual*, MySQL AB, Sweden, 2012.

Hindawi
Journal of Sensors
Volume 2018, Article ID 6167385, 10 pages
https://doi.org/10.1155/2018/6167385

Research Article

A Study of Prescriptive Analysis Framework for Human Care Services Based On CKAN Cloud

Jangwon Gim,[1] **Sukhoon Lee**,[1] **and Wonkyun Joo**[2]

[1]*Department of Software Convergence Engineering, Kunsan National University, Gunsan-si, Republic of Korea*
[2]*Biomedical HPC Technology Research Center, KISTI, Daejeon, Republic of Korea*

Correspondence should be addressed to Wonkyun Joo; joo@kisti.re.kr

Received 24 November 2017; Accepted 29 January 2018; Published 21 March 2018

Academic Editor: Ka L. Man

A number of sensor devices are widely distributed and used today owing to the accelerated development of IoT technology. In particular, this technological advancement has allowed users to carry IoT devices with more convenience and efficiency. Based on the IoT sensor data, studies are being actively carried out to recognize the current situation or to analyze and predict future events. However, research for existing smart healthcare services is focused on analyzing users' behavior from single sensor data and is also focused on analyzing and diagnosing the current situation of the users. Therefore, a method for effectively managing and integrating a large amount of IoT sensor data has become necessary, and a framework considering data interoperability has become necessary. In addition, an analysis framework is needed not only to provide the analysis of the users' environment and situation from the integrated data, but also to provide guide information to predict future events and to take appropriate action by users. In this paper, we propose a prescriptive analysis framework using a 5W1H method based on CKAN cloud. Through the CKAN cloud environment, IoT sensor data stored in individual CKANs can be integrated based on common concepts. As a result, it is possible to generate an integrated knowledge graph considering interoperability of data, and the underlying data is used as the base data for prescriptive analysis. In addition, the proposed prescriptive analysis framework can diagnose the situation of the users through analysis of user environment information and supports users' decision making by recommending the possible behavior according to the coming situation of the users. We have verified the applicability of the 5W1H prescriptive analysis framework based on the use case of collecting and analyzing data obtained from various IoT sensors.

1. Introduction

A number of sensor devices are widely distributed and used today owing to the accelerated development of IoT technology. In particular, this technological advancement has allowed users to carry IoT devices with more convenience, wear them for longer periods of time with greater comfort, and are evolving to become capable of collecting accurate user biometrics. As a result, there have been numerous studies in this area, such as those extracting the lifelog of users based on data collection from their various IoT devices and those detecting unusual signals from user biometrics. Among these, studies on smart healthcare services have been particularly active.

There are studies which conduct real-time monitoring and collection of user biometrics using wearable devices such as Fitbits and smartwatches to diagnose user behavior. Apple provides its "cloud" service in order to collect and use personal biosignal information communicated from an Apple watch. Fitbit, which contains a three-axis accelerometer, GPS, and other various sensors, is another device which also collects data and provides open API [1, 2]. Hence, users may utilize various devices based on their personal preference, despite all devices measuring the same bioinformation such as heartbeat, acceleration, and blood flow. It takes the advantage of choosing a suitable device according to the context and preference of a particular user; however, a solution is required to ensure data interoperability in integrating identical biometrics or data generated from different devices. Currently, studies are looking into user behavior analysis and prediction based on user biometrics generated based on the geographic information or atmospheric environment of a

specific area. The investigation aims at extracting the characteristics of users in the same circumstance and their unique patterns and also at analyzing their biometrics in order to provide users with customized healthcare services. Data for this investigation must not be device-dependent and should be collected and integrated according to each type of biometrics in order to provide corresponding smart healthcare services to specific user groups or individuals displaying similar patterns. Therefore, an integrated platform is required to collect, process, and manage data compatible with the cloud environment.

There are also studies aiming to analyze the behavioral patterns of specific users or groups of users based on user lifelog data and to predict behaviors based on the analysis. In particular, a previous study extracted elderly people's walking speed patterns and collected unusual signs such as reduced speed and frequent falls to predict incidents of falling or tripping. In other words, this study applied the data analysis methodology and attempted not only to collect past and present lifelog data, but also to develop techniques to predict possible future incidents. However, although this prediction technique warns users of eminent signs and potential future events based on their lifelog data, it does not provide information necessary for more effective decision making such as how to avoid such risks or events. As such, it is important to study methods that provide prescriptive analysis in preparation for specific signs and events.

Data analysis from a business aspect can be categorized into the following three categories [3]:

(i) Descriptive analytics: a set of technologies and processes to understand and analyze business outcomes using data.

(ii) Predictive analytics: prediction of results based on data input using a range of statistical and mathematical techniques. This method analyzes a variety of relationships based on predictions and suggests a model maximize business outcomes.

(iii) Prescriptive analytics: a set of mathematical techniques to determine complex targets, requirements, and limitations to improve business outcomes. This approach determines various alternatives and guides based on results drawn from descriptive analytics and predictive analytics.

Prescriptive analytics was launched as one of the hype cycle's emerging technologies announced by the Gartner Group in 2013 and has since received increasingly growing attention in business areas [4, 5]. Prescriptive analysis, as one type of data analysis technique, provides predictions and context-customized information. This technique is used to support more effective decision making based on various ideas when business decision makers, such as CTOs and CEOs, analyze and predict complex situations.

A previous study investigated big data analysis for healthcare in the big data environment [6]. Such analysis can provide smart healthcare services through event detecting mechanisms other than the existing data analysis methods. In other words, techniques for obtaining accurate healthcare data analysis in the big data environment, which allows the acquisition of a colossal amount of diverse data, play an instrumental role in providing smart human care services. Big data analysis is performed to support decision making based on insights into the user context and the planning, management, and learning for systematic data use. There are descriptive, predictive, and prescriptive analysis models used for data analysis. Each analysis model has been separately used according to existing services and purposes. However, it is necessary to merge them to draw final analytical results in the big data and cloud computing environments in which diverse data are integrated. Therefore, it is necessary to perform technological analysis on the users and their context to accurately diagnose their current status and also to integrate and analyze collected wearable sensor data and cognitive context information. In addition, various guides should be provided to users in urban environments where complicated and diverse events take place, in order to provide optimal smart healthcare services. Guides are well known to broaden the range of choices available for users in decision making. A prescriptive analysis technique should be introduced here to provide such guides in accordance with the user context [7]. To this end, this study proposes a 5W1H-based prescriptive analysis method based on the prescriptive analysis methodology to provide user-customized smart healthcare services and a data integration architecture featuring interoperability with data in cloud computing environments. The 5W1H approach can help to analyze contexts and situations of users because the 5W1H can seek to answer six basic questions such as "what," "when," "where," "who," "why," and "how" in gathering information about a plain. Therefore, nowadays, the 5W1H is being studied in the field of event processing to extract events from big data as a template or as a single ontology schema, to recognize the situation from events, and to detect abnormalities from events [8–12]. As one of the smart human care services, we apply the 5W1H approach to provide recommendation results for actions to be taken by users in specific situations.

Recently, there has been a pursuit of studies to reveal and increase the reusability of government-produced data, given the growing importance of data sharing in governmental and public organizations. This facilitated quality management of metadata and raw data, and several data portal systems supporting catalog management, and metadata interoperability were developed as a result. A representative example is the CKAN data portal. The CKAN is an open source data portal platform and is one of the tools used to build open data websites (https://ckan.org). The CKAN helps to manage data collection and sharing and is used by national and regional governments and various research institutions to collect a greater amount of data. In the CKAN system, users can save and share metadata schema to explain data catalogs and data resources saved on the system and as well as search for data. Diverse CKAN Extensions are being developed, and the system is applied in numerous fields for integrated management of open data. A typical example is the sharing of data produced in smart

cities. Maps and geographical information produced in smart cities can be saved and shared in the CKAN. Files of various types (XML, CSV, JSON, or SHP files) can be used in a RESTful way, and it is possible to connect the RDF store to save the ontology. The CKAN shares information on education, transportation infrastructures, energy consumption, local climate, and environment as an open data portal, thus allowing devices or persons to access various forms of data on services required by citizens or other forms of data. Furthermore, catalog and metadata saved on the CKAN can be harvested based on the OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting), allowing physical metadata repositories set up in other localities or separate CKANs used in other domains to be interconnected. In essence, it allows a cloud computing environment to be established. This indicates that it is possible to use the network as an interoperable base system for datasets defined by the different schemas. A previous study was conducted to save and analyze, in real time, data related to the economy, crime, and safety generated in smart cities based on the benefits of the CKAN including its functional extensibility and data interoperability [13, 14]. As a result of such efforts, there have been various systems capable of data integration and analysis using the CKAN. These features can also be applied in the integration of wearable sensor data. Data interoperability can be ensured by exchanging and harvesting metadata and schema required for the integration of raw data produced by wearable sensors, and sensor data can be incorporated based on integrated and interlinked schematic information.

The remainder of this study is structured as follows: Section 2 introduces related studies. Section 3 presents a smart healthcare framework for prescriptive analysis services. Section 4 explains user scenarios with sample data for explaining prescriptive analytics services based on 5W1H. Section 5 discusses the proposed framework. Section 6 presents conclusions and the future study directions.

2. Related Works

2.1. Sensor Information and Data Integration. A sensor is a module which observes stimuli and measures their value. This is called stimulus-sensor-observation (SSO) pattern, and such patterns are suitable for expressing and defining sensor information [15]. Semantic sensor network (SSN) ontology is an ontology developed by W3C based on SSO patterns [16]. Sensor information can be expressed in the existing sensor network environment as consisting of sensor nodes and sink nodes. In particular, W3C recently developed SOSA (sensor, observation, sample, and actuator) ontology to facilitate applications in the IoT environment. This is an extended version of SSN ontology and is capable of defining additional information such as sample, actuator, and provenance [17]. Sensor registry system (SRS) is a system for registering and managing sensor-related information [18]. The principal focus of this system is on metadata, and sensor information is registered in the heterogeneous ubiquitous sensor network environment. However, the SRS simply consists of metadata such as sensor location, sensor type, model name, and access type but does not collect the actual values registered by the sensor. SRS merely provides metadata from sensors to facilitate the integration of significant sensor information from specific sensor models. With this method, SRS is able to cover all sensors which are unable to transmit sensor data to cloud storage through the Internet. On the other hand, path prediction-based sensor registry system (PP-SRS) focuses on providing real-time metadata using mobile devices by applying path prediction technology to the SRS [19]. The existing SRS is able to provide sensor information as a location-based service as it manages sensor information including location coordinates. The PP-SRS provides sensor information by learning user paths through mobile devices such as users' smartphones and predicting the movement paths of users accordingly. This renders the PP-SRS more immediate and seamless.

2.2. Healthcare Platform and Analysis Method. Efforts to apply cloud computing to the healthcare field have been ongoing for several years now [20]. Doukas et al. attempted to access various healthcare information of patients with mobile devices through a cloud [21], and Rolim et al. suggested sensors attached to medical equipment in hospitals to save data in real time into a cloud service [22]. On the other hand, Zhang et al. proposed a healthcare cyber-physical system, aimed to primarily integrate medical data on a cloud environment—obtained from hospitals, the public health sector, research institutions, drug manufacturers, financial institutions, and wearable device manufacturers—and to analyze big data [23]. However, there are various issues in data sharing on cloud services [24]. In particular, patient information always bears security and privacy risks. Patient information saved on in-hospital systems such as medical devices or EMR contains protected health information. For this reason, the United States, for instance, enacted the HIPAA legislation and imposed strong regulations to control the storage of sensitive information on external servers. Contrarily, wide efforts are dedicated to converting patient healthcare data into a cloud service based on the recent advancements of digital healthcare technology, de-identification, and anonymization technology [25, 26]. In addition, the reliability of biosignal data collected from wearable devices has noticeably increased owing to the recent evolution of wearable devices such as smartwatches [27]. Diverse platforms have been developed to collect data, and it is now possible to measure patients' status and information on their surrounding environments using sensors through IoT and wearable devices [28]. Products such as S-Health by Samsung [29], Google Fit by Google [30], HealthKit by Apple [31], and Fitbit by Fitbit Inc. [32] all use their respective platforms to save data collected from smartwatches and various IoT devices onto a cloud server. Most of these platforms use RESTful API to provide users with collected data, and users can easily obtain their data through simple identification process and a URL. Owing to the evolution of wearable devices, there is a growing application of this technology to medical and healthcare studies using data obtained through wearable devices in place of medical devices [33]. Sensor technology coupled with wearable devices is already producing significant findings from studies,

and users can be provided with diverse healthcare services by integrating clinical data obtained from sensors on a cloud. In order to achieve this, a cloud-based framework is required for medical information integration. In addition, a prescriptive analytics service should be in place for accurate analysis and prediction of user context in order to provide custom information and to support users' decision making.

In relation to analytic services, several studies have been launched in order to provide context-customized services by analyzing data generated from humans to support human care services. For example, there has been research conducted on tools, technology, and methodology for healthcare data value analysis [34]. The prescriptive analysis technique in the medical support area allows the provision of more varied information when used for treating patients than for analyzing users' current contexts and is used to benchmark the resulting rise in hospital revenue [35]. In other words, various guides have been suggested to support family doctors to treat patients, providing predicted outcomes according to the particular treatment and thus allowing doctors to choose the most suitable guide. As a result, hospitals are able to analyze outcomes based on the chosen guide. This prescriptive analysis technique is also used in business situations to compare performance outcomes. Thus, it is necessary to apply the prescriptive analysis technique to users and patients. Prescriptive results produced through prescriptive analysis can be interlinked with healthcare services. Further studies are required on prescriptive analysis techniques which can provide more appropriate results through analysis and prediction of user context information.

3. Architecture

This chapter explores data sharing platforms for providing prescriptive analysis-based human care services.

It proposes the implementation and design of cloud-based data analysis service and elaborates on a set of processes used for human care services. The major principle of the proposed architecture design is to support appropriate decision making in decision-requiring situations such as disasters, activities, and events based on user context information. The proposed system includes the 5W1H (what, when, where, who, why, and how) categorization process to perform a prescriptive analysis. The proposed architecture diagnoses the current contexts of users based on diagnostic data saved on the CKAN-based cloud computing environment and also predicts eminent situations, events, and incidents. It provides users with prescriptive guides on when and how they must respond to a specific event based on predicted results. Figure 1 presents the system architecture for data integration using a cloud environment and for providing prescriptive analysis services. Therefore, the architecture is divided into two layers. The first is the cloud computing layer, and the second is the layer for data analysis. The cloud computing layer is shown at the bottom of Figure 1. The cloud computing environment is indispensable regarding the sharing of a large amount of big data. Hence, the cloud computing environment enables the integration of data and the sharing of analytical tools. Also, open data used in different application domains (environment, economy, and disasters) can be interlinked using the CKAN to establish a cloud computing environment. The data obtained from wearable sensors are saved on a repository where sensor schema is defined. Context information in a cloud computing environment is established through wearable sensors (temperature, humidity, carbon dioxide, and heartbeat) and various open data. Here, the reliability and accuracy of wearable sensor data must be guaranteed for more accurate context analysis, and the transparency must also be guaranteed in data integration. Upon the diversification of wearable sensor types, however, problems may arise in data integration when the data storing schema do not correspond even though the sensor data are measured in an identical environment as the different types of products used for measurement or when no common data are in place that can interlink separate datasets. Therefore, catalogs should be shared which contain standardized metadata for data integration. The catalogs and their properties and types are shared on the CKAN data portal. The CKAN data portal and repositories in each domain can mutually harvest and share information based on the OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting). This extends CKAN harvester to parse OAI-PMH metadata sources and import datasets. It supports metadata schemas such as DC (Dublin Core) and CMDI (Component MetaData Infrastructure). Through the OAI-PMH protocol, CKANs in the cloud can share their catalogs and datasets, and the metadata and properties included in each catalog are integrated and extended based on the existing registered common concepts in the cloud. Therefore, common concepts are continually extended by newly added metadata and attributes. As a result, it is used as a knowledge graph for inference of implicit situation information, derivation of relationships between sensor data, and analysis of sensor data. Therefore, the context information of each user collected from the wearable sensors/devices is transmitted to each CKAN server, and each transmitted sensor data is integrated and stored in CKAN cloud. When the data is integrated into the CKAN cloud, catalogs, ontologies, and common concepts can be applied to integrate each sensor data from different sensors. The second data analysis layer is shown at the top of Figure 1. The data analysis layer consists of three components as shown in Figure 1 (2-1, 2-2, 2-3). The first is to extract events from the aggregated data obtained from the cloud computing layer, where the events are classified as 5W1H. We use three analytical techniques: descriptive analytics, predictive analytics, and prescriptive analytics (2-1). In the 2-1 step, events are extracted based on 5W1H factors with rules and conditions which can explain users' situations and contexts. At the first step of 2-1, descriptive analytics method extracts basic statistics information from the knowledge base stored in the CKAN cloud. After that, based on the results of 2-1 phase, future events coming up to the user (e.g., increase heartbeat, lower body temperature, and a surge of carbon dioxide) are extracted (2-2). The following stage predicts trends using machine learning algorithm based on the extracted context information. Potential variables, conditions, and limitations in specific contexts are

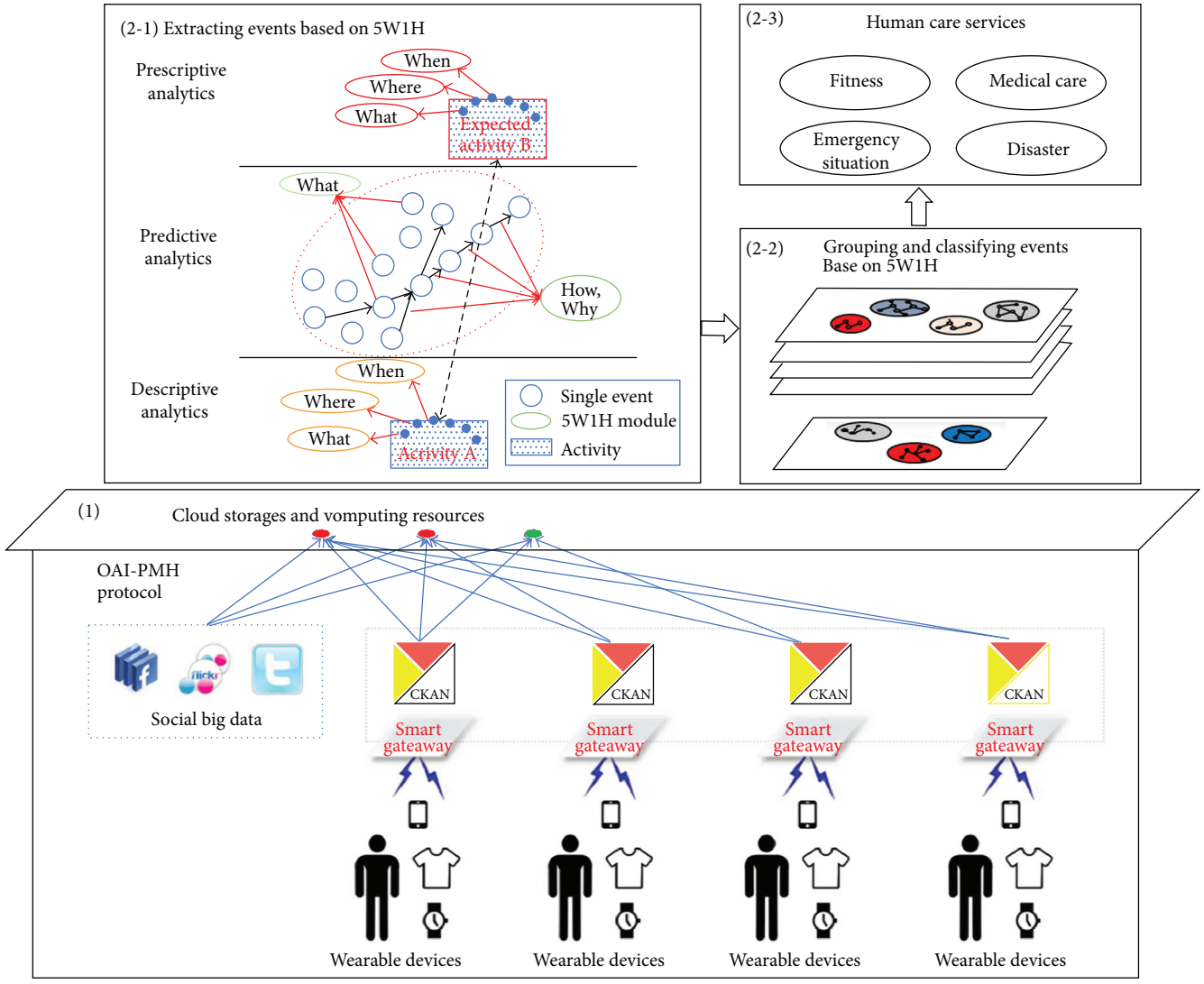


Figure 1: Overall architecture based on CKAN.

selected either through supervised learning or through unsupervised learning. In the 2-3 phase, the diagnosed context and predicted outcomes obtained from these results are used to provide guides which help users to make decisions, and this process is defined as prescriptive analysis. Prescriptive information can be obtained through this prescriptive analysis, which was not possible through diagnosis and predictions. Users can make prompt and accurate decisions from a group of available response options. In this way, smart human care services can be provided to users who may have to deal with big data and various other situations.

4. Service Scenario

This chapter explains how datasets are collected and the implementation of prescriptive analysis with wearable devices using the CKAN through an overall scenario. The CKAN-based data integration method and the prescriptive analysis process are presented through a specific case. The case describes how firefighters collect integrated data through turnout gears and available devices and how a human care service is provided based on this mechanism.

4.1. Overall Scenario. Figure 2 shows the flow and structure of the service scenario. This scenario starts from the wearable devices of the user. The user is wearing a smartphone, smartwatch, and smart clothes. In the CKAN-based data integration area, (1) data are extracted from the user's wearable devices, and the data from the wearable devices are integrated using ontology such as SSN ontology. (2) Extracted sensor values are processed into datasets as a unified schema for the same sensor information by matching this information with SSN ontology. (3) The datasets are stored in the CKAN repository. On the other hand, prescriptive analysis (PA) should first be conducted in order to provide human care services as suggested by this paper. For this, (4) the dataset to be defined and its corresponding ontology is to be searched. Searched datasets are, respectively, processed through descriptive analytics, predictive analytics, and prescriptive analytics. (5) Scripts resulting from these analyses are stored

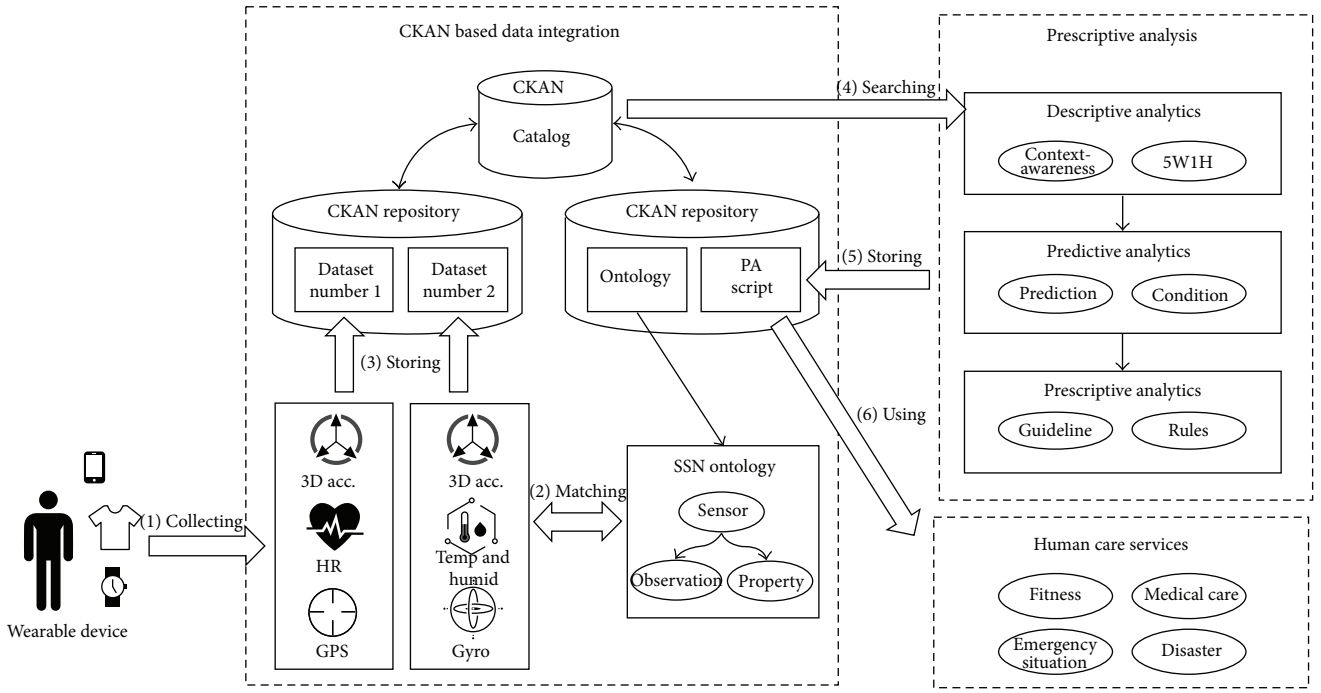


Figure 2: A conceptual model for scenario.

Table 1: Collected sensor types of smart garments.

Sensor	Measurement type	Type name	Datatype	Unit of measure
External temperature sensor	Temperature	ex_temp	Double	Celsius, Fahrenheit
Heat flux sensor	Temperature	hf_temp	Double	Celsius, Fahrenheit
GPS	Latitude	lat	Double	E, W, S, N
	Longitude	lon	Double	
	Altitude	alt	Double	
3D accelerometer	x-axis acceleration	acc_x	Double	mV/g
	y-axis acceleration	acc_y	Double	
	z-axis acceleration	acc_z	Double	
Gas sensor	CO concentration	co	Double	ppm
	CO_2 concentration	co2	Double	
HR sensor	Heart rate	hr	Integer	bpm
SPO_2 sensor	Arterial oxygen saturation	spo2	Integer	Percentage

again on the CKAN repository as PA scripts. (6) Diverse human care services are provided using PA scripts afterward.

4.2. CKAN-Based Data Integration. For a more expansive scenario, it must first be assumed that the firefighters use wearable devices including sensors attached to their turnout gear and a smartwatch. The ProeTEX project, for instance, developed smart garments such as turnout gear, t-shirt, and a pair of boots with 12 sensors, 2 alarms, and 4 network modules [36]. Table 1 provides the definition of different sensor types used in this scenario out of the group of sensors attached to smart garments. Based on the sensors, we extracted features such as type, type name, datatype, and unit of measure. The features can be important elements to integrate sensor data considering their semantics from diverse IoT sensors when the data derived from IoT sensors is integrated based on common concepts such as an ontology schema stored in the CKAN cloud.

Figure 3 demonstrates how the external temperature sensor is associated to SSN ontology among the sensors on the smart garments. The sensor, measurement type, and type name columns in Table 1 are mapped into "sosa: Sensor," "sosa: ObservableProperty," and "sosa: Observation," respectively. Sensor observations generated from a single user can be bound through "sosa: FeatureofInterest" like the case of "ProeTEX/user1." Various wearable device sensors and lifelogs can be integrated at the SSN ontology schema level through ontology matching. RDF generated as a result of matching is stored on the CKAN repository with corresponding datasets. In fact, SOSA is capable of a raw data

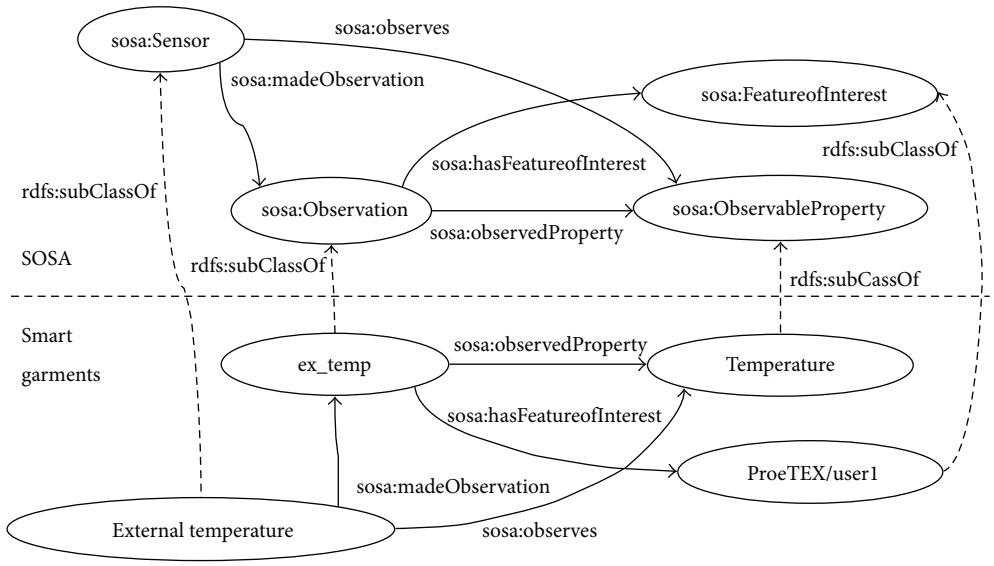


Figure 3: A sensor ontology to merge ontology instances.

Table 2: Features classified as 5W1H.

Features	Measurement type	5W1H
ex_temp	Temperature	What
hf_temp	Temperature	What
lat, lon, alt	Latitude, longitude, altitude	Where
acc_x, acc_y, acc_z	*x*-, *y*-, *z*-axis acceleration	What
co	CO concentration	What
co2	CO2 concentration	What
hr	Heart rate	What
spo2	Arterial oxygen saturation	What
o2	Oxygen ratio of firefighter respirator	What
user_id	User identifier	Who
time	Timestamp	When
date	Timestamp	When

(observation value) level expression, but a schema level integration was considered in this scenario.

4.3. 5W1H-Based Prescriptive Analysis. The prescriptive analysis method suggested in this paper begins by searching the data stored on the CKAN. As shown in Figure 2, descriptive analytics, predictive analytics, and prescriptive analytics are consecutively conducted in a prescriptive analysis. Here, data required for descriptive analytics are categorized into 5W1H using context-awareness technology. Table 2 shows how the features of each sensor are classified into 5W1H through descriptive analytics.

The 5W1H classification is performed depending on the "sosa:observableProperty" of each measurement type during the ontology matching described in Section 4.2. In the step of a predictive analysis, each feature can be extracted and classified through machine learning algorithms and data mining methods considering the 5W1H. First, data collected from the sensors are learned, and a concrete phenomenon is predicted. Prediction results and conditions are categorized into factors and defined based on the predicted situation. For example, a phenomenon is predicted using many features such as "the external temperature will rise by about 10 degrees in 10 minutes." Then, features such as "in 10 minutes" and "the external temperature will rise by about 10 degrees" and their predicted values are extracted. Moreover, situations caused by specific conditions can be predicted as well. It is possible to suggest a prediction result such as "it is dangerous if the external temperature is higher than 200 degrees Celsius." Through this process, conditions like "the external temperature is higher than 200 degrees Celsius" can be extracted. Table 3 shows results of a concrete situation and extractable conditions through predictive analytics.

In summary, prescriptive analytics suggests a guideline for specific situations to users. Users are vulnerable to various risks when they do not have any previous experience or are unfamiliar with the situation because they are unsure on what to do. The application of the previously defined predictive situation and limitation will allow suggesting a systematically more immediate and effective guideline to users. Table 4 presents guidelines and rules drawn from the combinations of predictions and conditions.

Prescriptive analysis is highly effective, particularly in critical firefighting situations. The example scenario suggests a set of guidelines to help firefighters predict and avoid danger when they are in, or are highly likely to be in, a dangerous situation. No matter how experienced a particular firefighter may be, these kinds of systematic context-awareness guidelines are highly useful when they find themselves in extreme fire incidents [37].

5. Discussion

Contexts should be accurately diagnosed and analyzed to provide prescriptive analysis services for diverse situations

Table 3: Predictions and conditions extracted from predictive situations and limitations using 5W1H factors.

Predictive situations (PS), limitations (L), and 5W1H features	Predictions (p) and conditions (c)
PS1: the external temperature will rise by about 10 degrees in 10 minutes. 5W1H ➔ {What, When, How}	p1) time: after 10 min p2) ex_temp: 10°C increase
PS2: the heart rate will rise by 40 bpm in 10 minutes, and oxygen ratio of the respirator falls by 10%. 5W1H ➔ {What, When, How}	p1) time: after 10 min p3) hr: 40 bpm increase p4) o2: 10% decrease
CO2 concentration will rise by 1000 ppm in 5 minutes. 5W1H ➔ {What, When, How}	p5) time: after 5 min p6) co2: 1000 ppm increase
It is dangerous if the external temperature is higher than 200 degrees. 5W1H ➔ {What, How}	c1) ex_temp: more than 200°C
It is dangerous if the oxygen ratio of the respirator is lower than 10%. 5W1H ➔ {What, How}	c2) o2: less than 10%
It is dangerous if the CO2 concentration is higher than 10,000 ppm. 5W1H ➔ {What, How}	c3) co2: more than 10,000 ppm

Table 4: Guidelines and rules with 5W1H factors.

Prescriptive rule	Guideline (GL)
{p1, p2, p3, c1} → GL1	GL1) Go outside within 5 minutes. 5W1H ➔ {What, When, Where}
{p1, p3, p4, c2} → GL2	GL2) Start charging your firefighter respirator within 5 minutes. 5W1H ➔ {What, When, How}
{p5, p6, c3} → GL3 ∨ GL4	GL3) Put on your firefighter respirator immediately. 5W1H ➔ {What, When} GL4) Go outside immediately. 5W1H ➔ {What, When}

in a cloud computing environment. Data collected for this purpose should be managed transparently and quickly integrated. Semantic interoperability between wearable sensor data should also be guaranteed. To achieve this, a CKAN-based cloud data portal which is able to share common schema or catalogs is proposed. In order to perform prescriptive analysis by drawing analytical results from the data portal, the following factors should be considered:

(i) The need for machine-readable raw data

(ii) The need for data interlinking, to link the meanings of data

(iii) The need for sharing metadata and catalogs used on different domains

(iv) The need for stable processing and complex event processing techniques for big sensor data

Furthermore, in order to provide various guidelines based on prescriptive analytics results excluding those from descriptive and predictive analytics, a set of rules and limitations should be established at the prescriptive analysis stage. To achieve this, simulations of diverse situations should be performed to generate multiple candidate recommendations for suggesting an optimum guide by configuring environmental variables and conditions.

With the increase of large-scale data and the innovative development of data processing technology, an environment that can handle general knowledge, as well as specific areas, is being created. For example, Graph search on Facebook has suggested a new search method that can find the most relevant "people," "photos," "places," and "interests" on Facebook. This shows that semantic retrieval is possible over existing syntax search through graphical knowledge structure. Google's knowledge base also builds a database of vast amounts of knowledge and common sense and uses it for a variety of intelligent services. Therefore, this knowledge base can be usefully used in the field of smart human care service. Based on the knowledge base, we can derive various smart services considering user environment or find additional semantic information which cannot be derived from existing single or partial sensor data by using various sensor data. For this purpose, the proposed CKAN cloud environment that can share common concept as one of knowledge base and can collect and store big data will be useful. Specific situations can be simulated through an accumulation of scenarios, rules, and limitations for each situation, and this increases the accuracy of prescriptive analysis services. Interlinking information from many different domains saved on the cloud generates situation perceiving knowledge graphs, which allows the derivation of new rules from a set of rules and limitations. In this manner, the range of guidelines provided by prescriptive analysis services can continue to expand, and it will be possible not only to provide information on specific singular situations but also to suggest guidelines for possible situations in complex events and areas of interest. Hence, the 5W1H method enables efficient extraction of events that are meaningful to the user from a vast amount of events. In other words, when a dangerous disaster situation or difficult decision-making situation is encountered from the viewpoint of a specific user, an action guide can be provided to the user through an indicator analysis. In addition, when real-time collection and processing is required, a template of a formalized form is required to automatically classify and extract meaningful data within a fixed time, and the proposed 5W1H method can be applied.

6. Conclusion

The wearable sensor technology is advancing at a remarkable pace, along with the recent development of IoT technology. As a result, sensors have generated countless data, and current studies are investigating methods to draw meaningful analysis from big data. Various studies are currently being conducted in the human care services area to diagnose and analyze user contexts in various situations. However, such analysis services do not provide guidelines on how to respond to eminent risk factors or events. In other words, they only provide analytical results to users and do not offer any guidelines as to when a user should take action, what kind of action they should take, and how to perform such action in the event of a dangerous situation (fall, car accident, fire, and disaster). In order to analyze and diagnose given situations accurately, it is imperative to integrate data in many different forms that are publicly shared by many local and related fields and as well as the user's wearable sensor information. An analysis technique is also required to provide an appropriate prescriptive analysis result (guideline) based on integrated data. This paper proposed an architecture to perform prescriptive analysis based on CKAN, which is widely used for data sharing to solve such issues. Using a 5W1H-based prescriptive analysis method presented in this study, the applicability of the architecture was verified based on a user scenario. The proposed system architecture can be widely applied in many fields for analyzing smart cities and urban environments and responding to situations. In this way, smart human care services can be provided to users wearing wearable sensors and devices.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Acknowledgments

This work was supported by the "Development of biomedical data network analysis technology for dementia researches (K-18-L03-C02)" funded by Korea Institute of Science and Technology Information. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2017R1C1B2012065).

References

[1] T. Fritz, E. M. Huang, G. C. Murphy, and T. Zimmermann, "Persuasive technology in the real world: a study of long-term use of activity sensing devices for fitness," in *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems - CHI '14*, pp. 487–496, Toronto, ON, Canada, 2014, ACM.

[2] T. Rault, A. Bouabdallah, Y. Challal, and F. Marin, "A survey of energy-efficient context recognition systems using wearable sensors for healthcare applications," *Pervasive and Mobile Computing*, vol. 37, pp. 23–44, 2017.

[3] S. K. Song, D. H. Jeong, J. Kim, M. Hwang, J. Gim, and H. Jung, "Research advising system based on prescriptive analytics," in *Future Information Technology*, pp. 569–574, Springer, Berlin, Heidelberg, 2014.

[4] A. Linden and J. Fenn, "Understanding Gartner's hype cycles," Strategic Analysis Report, R-20-1971. Gartner, Inc., Stamford, CT, USA, 2003.

[5] S. John Walker, "Big data: a revolution that will transform how we live, work, and think," *International Journal of Advertising*, vol. 33, no. 1, pp. 181–183, 2014.

[6] A. Belle, R. Thiagarajan, S. M. R. Soroushmehr, F. Navidi, D. A. Beard, and K. Najarian, "Big data analytics in healthcare," *BioMed Research International*, vol. 2015, Article ID 370194, 16 pages, 2015.

[7] W. Raghupathi and V. Raghupathi, "An overview of health analytics," *Journal of Health & Medical Informatics*, vol. 4, p. 132, 2013.

[8] S. Sharma, R. Kumar, P. Bhadana, and S. Gupta, "News event extraction using 5W1H approach & its analysis," *International Journal of Scientific & Engineering Research*, vol. 4, no. 5, pp. 2064–2068, 2013.

[9] L. Xie, H. Sundaram, and M. Campbell, "Event mining in multimedia streams," *Proceedings of the IEEE*, vol. 96, no. 4, pp. 623–647, 2008.

[10] W. Wang, D. Zhao, L. Zou, D. Wang, and W. Zheng, "Extracting 5W1H event semantic elements from Chinese online news," in *International Conference on Web-Age Information Management*, pp. 644–655, Springer, Berlin, Heidelberg, 2010.

[11] W. Wang, "Chinese news event 5W1H semantic elements extraction for event ontology population," in *Proceedings of the 21st International Conference Companion on World Wide Web - WWW '12 Companion*, pp. 197–202, Lyon, France, 2012, ACM.

[12] W. Zhou, Y. Zhang, X. Su, Y. Li, and Z. Liu, "Semantic role labeling based event argument identification," *International Journal of Database Theory and Application*, vol. 9, no. 6, pp. 93–102, 2016.

[13] T. Käfer, J. Umbrich, and A. Hogan, "Polleres: Towards a Dynamic Linked Data Observatory," in *Workshop on Linked Data on the Web(LDOW)*, vol. 937, 2012.

[14] S. Lalithsena, P. Hitzler, A. Sheth, and P. Jain, "Automatic domain identification for linked open data," in *2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*, vol. 1, pp. 205–212, Atlanta, GA, USA, November 2013, IEEE.

[15] K. Janowicz and M. Compton, "The stimulus-sensor-observation ontology design pattern and its integration into the semantic sensor network ontology," in *Proceedings of the 3rd International Conference on Semantic Sensor Networks SSN'10*, vol. 668, pp. 64–78, Shanghai, China, November 2010, CEUR-WS.org.

[16] M. Compton, P. Barnaghi, L. Bermudez et al., "The SSN ontology of the W3C semantic sensor network incubator group," *Web Semantics: Science, Services and Agents on the World Wide Web*, vol. 17, pp. 25–32, 2012.

[17] Semantic Sensor Network Ontology, http://www.w3.org/TR/vocab-ssn/.

[18] D. Jeong, "Framework for seamless interpretation of semantics in heterogeneous ubiquitous sensor networks," *International Journal of Software Engineering and Its Applications*, vol. 6, no. 3, pp. 9–16, 2012.

[19] S. Lee, D. Jeong, D. K. Baik, and D. K. Kim, "Path prediction method for effective sensor filtering in sensor registry system," *International Journal of Distributed Sensor Networks*, vol. 11, no. 7, article 613473, 2015.

[20] A. M. H. Kuo, "Opportunities and challenges of cloud computing to improve health care services," *Journal of Medical Internet Research*, vol. 13, no. 3, article e67, 2011.

[21] C. Doukas, T. Pliakas, and I. Maglogiannis, "Mobile healthcare information management utilizing cloud computing and android OS," in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, pp. 1037–1040, Buenos Aires, Argentina, August 2010, IEEE.

[22] C. O. Rolim, F. L. Koch, C. B. Westphall, J. Werner, A. Fracalossi, and G. S. Salvador, "A cloud computing solution for patient's data collection in health care institutions," in *2010 Second International Conference on eHealth, Telemedicine, and Social Medicine*, pp. 95–99, St. Maarten, Netherlands, February 2010, IEEE.

[23] Y. Zhang, M. Qiu, C. W. Tsai, M. M. Hassan, and A. Alamri, "Health-CPS: healthcare cyber-physical system assisted by cloud and big data," *IEEE Systems Journal*, vol. 11, no. 1, pp. 88–95, 2017.

[24] S. P. Ahuja, S. Mani, and J. Zambrano, "A survey of the state of cloud computing in healthcare," *Network and Communication Technologies*, vol. 1, no. 2, 2012.

[25] L. Devadass, S. S. Sekaran, and R. Thinakaran, "Cloud computing in healthcare," *International Journal of Students' Research in Technology & Management*, vol. 5, no. 1, pp. 25–31, 2017.

[26] N. Mohammed, B. C. M. Fung, P. C. K. Hung, and C.-K. Lee, "Centralized and distributed anonymization for high-dimensional healthcare data," *ACM Transactions on Knowledge Discovery from Data*, vol. 4, no. 4, pp. 1–33, 2010.

[27] J. Bughin, M. Chui, and J. Manyika, "Clouds, big data, and smart assets: ten tech-enabled business trends to watch," *McKinsey Quarterly*, vol. 56, no. 1, pp. 75–86, 2010.

[28] F. De Arriba-Pérez, M. Caeiro-Rodríguez, and J. Santos-Gago, "Collection and processing of data from wrist wearable devices in heterogeneous and multiple-user scenarios," *Sensors*, vol. 16, no. 9, p. 1538, 2016.

[29] Samsung Health, https://health.apps.samsung.com.

[30] Google Fit, https://www.google.com/fit.

[31] Apple HealthKit, https://www.apple.com/ios/health.

[32] Fitbit, https://www.fitbit.com.

[33] B. Reeder and A. David, "Health at hand: a systematic review of smart watch uses for health and wellness," *Journal of Biomedical Informatics*, vol. 63, pp. 269–276, 2016.

[34] T. L. Strome and A. Liefer, *Healthcare Analytics for Quality and Performance Improvement*, Wiley, Hoboken, NJ, USA, 2013.

[35] D. B. Nash, "Harnessing the power of big data in healthcare," *American Health & Drug Benefits*, vol. 7, no. 2, pp. 69-70, 2014.

[36] D. Curone, E. L. Secco, A. Tognetti et al., "Smart garments for emergency operators: the ProeTEX project," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 3, pp. 694–701, 2010.

[37] S. Patel, H. Park, P. Bonato, L. Chan, and M. Rodgers, "A review of wearable sensors and systems with application in rehabilitation," *Journal of NeuroEngineering and Rehabilitation*, vol. 9, no. 1, p. 21, 2012.

Hindawi
Journal of Sensors
Volume 2018, Article ID 9845321, 10 pages
https://doi.org/10.1155/2018/9845321

Review Article

Theory and Application of Audio-Based Assessment of Cough

Yan Shi, He Liu, Yixuan Wang, Maolin Cai, and Weiqing Xu

School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China

Correspondence should be addressed to Yan Shi; shiyan@buaa.edu.cn and Weiqing Xu; weiqing.xu@buaa.edu.cn

Received 17 August 2017; Accepted 27 November 2017; Published 6 March 2018

Academic Editor: Mucheol Kim

Cough is a common symptom of many respiratory diseases. Many medical literatures underline that a system for the automatic, objective, and reliable detection of cough events is important and very promising to detect pathology severity in chronic cough disease. In order to track the development status of an audio-based cough monitoring system, we briefly described the history of objective cough detection and then illustrated the cough sound generating principle. The probable endpoints of cough clinical studies, including cough frequency, intensity of coughing, and acoustic properties of cough sound, were analyzed in this paper. Finally, we introduce some successful cough monitoring equipment and their recognition algorithm in detail. It can be obtained that, firstly, acoustic variability of cough sounds within and between individuals makes it difficult to assess the intensity of coughing. Furthermore, now great progress in audio-based cough detection is being made. Moreover, accurate portable objective monitoring systems will be available and widely used in home care and clinical trials in the near future.

1. Introduction

Cough is a common but complicated symptom of respiratory diseases. This symptom is also the very reason why people seek medical advice in America and China [1, 2]. Even though the importance of cough diagnosis is well admitted by academic organizations [2–4] in recent years, there is no gold standard to assess cough due to the lack of objective and accurate measures of cough frequency and severity [5]. When cough becomes chronic, it is so extremely unpleasant and distressing that the life quality of chronic cough patients has significant reductions [6]. The health care cost, medical consultations, and medication use hence become a heavy burden for them [7]. The assessment of cough severity is subjective at present: it contains visual analogue scales (VAS), health-related quality of life (HRQOL), Leicester cough questionnaire (LCQ), cough-specific quality of life questionnaire (CQLQ), and so on [8, 9]. They have been validated in chronic or acute cough in clinical trials [9]. However, these tools are completed either by the patient himself or by a parent [10], and hence it is in conflict with the standard that the primary outcome measure of clinical trials should be objective. Moreover, some literatures have shown that the objective cough frequency monitors may not have much to do with the subjective assessment methods of cough [11, 12]. It may be due to the different standards used in these tools. The medical literature indicates the necessity of an objective and reliable tool to measure the severity of cough [3].

As early as the 1950s, some attempts to monitor cough objectively have been made [13]. Since then, there have been three main ways to record cough. One is based on the airflow measurement at the mouth to obtain the flow dynamics of cough [14, 15]. However, this method is not suitable for continuous monitoring in the outpatient environment [16]. The second is based on the movement of the chest. For example, some researchers [17] invented an accelerometer-based system that used an accelerometer placed at the volunteer's chest wall to record cough events, but such system required researchers to count coughs manually. The measurement of cough sounds, the last one, has been more universal because of advances in computer technology and the availability of portable digital sound recording devices [18].

The underlying disease determines the physical character of the cough sound [19]; cough has been described as dry, wet, loose, or whooping, depending on the amount of

expectoration and sound quality. Therefore, methods based on cough sounds for counting and classifying cough events have been developed. This article focuses on audio-based methods and systems for the analysis and measurement of cough.

2. Cough Sound Basics

An airflow and acoustic signal for a cough is seen in Figure 1. As shown in the figure, a deep inspiration usually starts a classical cough, followed by glottis closure. During the glottis closure, respiratory muscles contract against the closed glottis and then the sudden opening of the glottis occurs with transient and fast expiratory airflow accompanied by the typical cough sound. Sometimes, several further partial glottis closures lead to some extra voiced sounds, which also called a cough sequence [6, 20]. However, the origin of cough sounds is still unclear because laryngeal structures and the resonance of the nasal and thoracic cavity are all involved in cough and their roles in cough are uncertain to some extent [21].

The typical cough sound is usually divided into three phases (in Figure 2) [3]: (1) an explosive expiration due to the glottis suddenly opening, (2) the intermediate phase with the attenuation of cough sounds, and (3) the voiced phase due to the closing of the vocal cord. In fact, there are a variety of patterns of cough that occur; for example, some cough sounds only have two phases (the intermediate phase and the voiced phase) and the explosive phase usually prolonged because of some diseases.

3. Endpoints of Objective Cough Assessment

Cough frequency evaluation is considered to be a gold standard for the objective evaluation of cough [8]. Besides it, the intensity of coughing, the pattern of coughing, and the acoustic properties of cough sounds may be used as clinical endpoints [3, 8].

3.1. Cough Frequency. According to the second part, even if there exist a variety of patterns of coughing making it difficult to identify and quantify cough, coughing can be quantified in four different ways [22]:

(1) Explosive cough sounds: the number of characteristic explosive cough impulses

(2) Cough seconds: the quantity of seconds and hours having at lowest an explosive phase

(3) Cough breaths: respiration rates including at least a cough

(4) Cough epochs: the number of cough sounds with no more than two seconds of each coughing interval.

The effectiveness of any of these metrics over the other is still ongoing research.

These methods are used for counting cough events, and it is not clear whether any of them is more valid than the other. There is a good linear relationship between explosive cough sounds and cough seconds under different circumstances. Cough epochs are less related to explosive cough sounds [6, 23]. As a result, current chronic cough frequency monitors usually use explosive cough sounds to evaluate cough frequency [22].

24-hour cough frequency is proved correct and effective in a longitudinal observational study of 33 healthy subjects with acute cough [24]. 4-hour cough frequency is found to be responsive to improvements in cough severity following trials of therapy in 100 patients with chronic cough [25].

3.2. Cough Intensity. Chronic cough is a common condition related to significant physical and psychological morbidity. But there is a weak relationship between health-related quality of life and cough frequency [26]. This has resulted in cough intensity of some patients that may be significant [27, 28]. The intensity of voluntary, induced, and spontaneous cough has been researched [29, 30], and peak cough flow rate, oesophageal pressure, and gastric pressure are important and relevant measures of cough intensity in patients with chronic cough [31]. Limitations of above indexes are that they are either invasive or impractical to measure in an ambulatory setting [8]. Therefore, cough sound is a potential measure of the intensity of coughs.

Cough intensity may be measured by cough sound power, peak energy, and mean energy [31, 32], and these indexes can be calculated for a time window with a duration of 0.5 second from one set of phase 1, the explosive phase of cough sound [32]. However, more researches should assess cough sound responsiveness as a measure of cough intensity.

3.3. Cough Patterns and Acoustic Properties of Cough Sounds. Cough patterns and the quality of a patient's cough sound may reflect useful information about their condition. Cough patterns and some cough sound features may be endpoints of clinical experiments [33]. However, more researches should be undertaken to study the relationship between cough patterns (or acoustic properties of cough sounds) and the illness-triggering cough.

4. Audio-Based Cough Monitoring Systems

In the analysis of cough sounds, researchers focus on two aspects: one is the study of cough recording and monitoring equipment and the other is the study of a cough sound processing algorithm.

As previously mentioned, cough frequency is the most valuable index of objective cough assessment [34]. Meanwhile, cough frequency monitoring technology is also the most mature one among the present objective cough assessment technology researches [22]. Manual counting of cough sounds remains the reference standard because compared with other tools, the human ear performs best in counting cough events [35, 36]. Even so, the arduous and time-consuming nature of manually counting cough events restricts its feasibility in larger-scale studies and clinical application [35]. Therefore, automatic monitoring of cough sounds is a development trend of objective cough evaluation.

Currently, there is no accurate fully automated cough detection system available, because it is challenging to

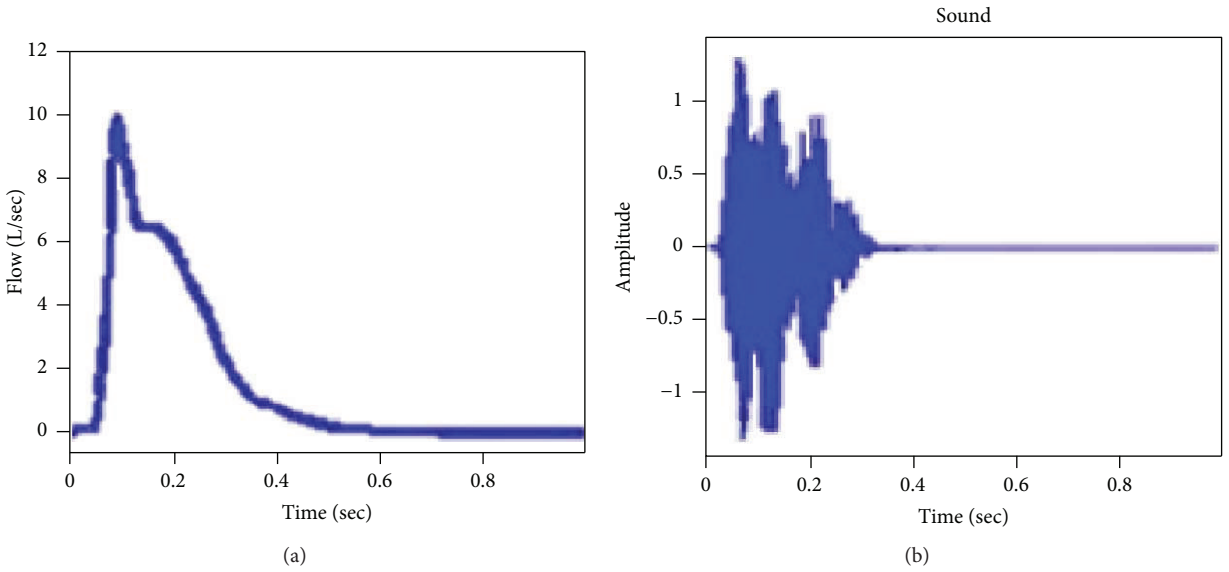


Figure 1: An airflow (a) and acoustic signal (b) for a cough. The airflow has been measured at the mouth and the cough sound at the sternal manubrium of the patient.

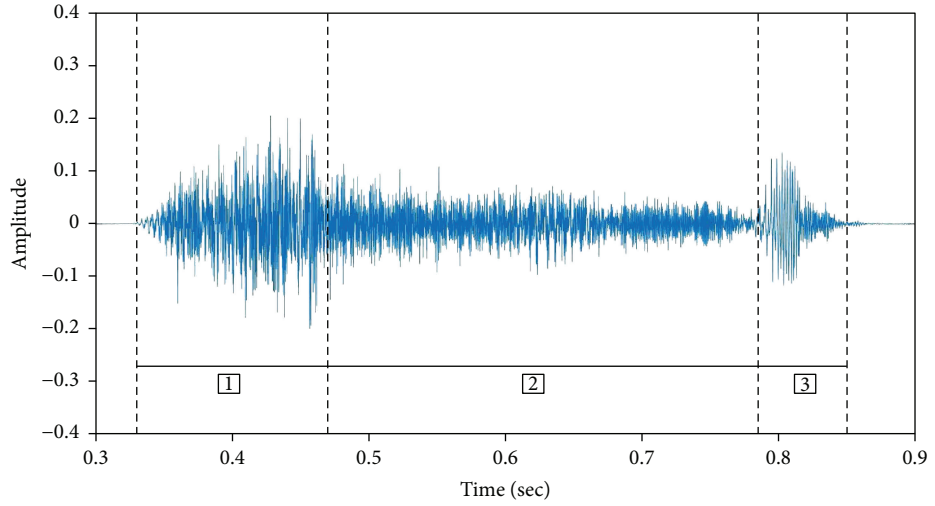


Figure 2: Typical cough sound (1: explosive phase, 2: intermediate phase, and 3: voiced phase).

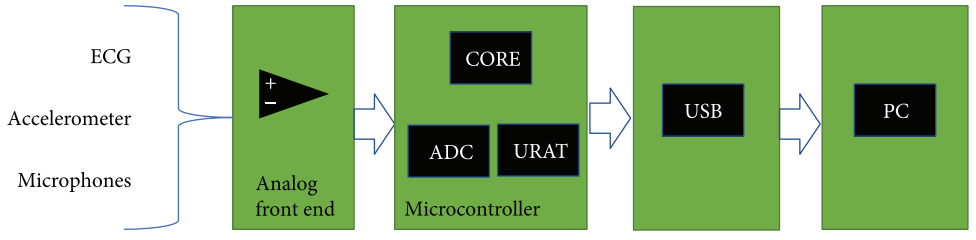


Figure 3: Block diagram of the acquisition system.

replicate the performance of the human ear to detect cough sounds. Recently, technological advances in digital storage devices and sound sensors make it portable and accurate to record cough sounds. Several cough monitors have been developed, and they adopt audio signals alone (microphone and/or contact microphone) or in combination with other sensors such as accelerometers, pneumographic belt, electromyography electrodes, electrocardiography electrodes, induction plethysmography, and pulse oximeter (in Figure 3). Drugman et al. [37] found that the audio

microphone performed best among these sensors for cough detection. Thus, we can divide these existing cough frequency monitors into two sections: one only uses audio signal and the other uses mixed signals.

4.1. Audio-Only Cough Monitors. Some cough monitors use the audio only. The Hull Automatic Cough Counter (HACC, Castlefield Hospital, Hull, UK) uses a free-field microphone to record ambulatory sound around 24 h (Figure 4(a)). It uses an artificial neural network (ANN) to detect coughs after signal processing. The system can label coughs automatically but count coughs manually. In a test of 33 patients with chronic cough, the HACC presents a sensitivity of 80% (range from 55 to 100%) and a specificity of 96% [38]. Over 24 h recording and further assessment in different conditions are required.

The Leicester Cough Monitor (LCM, Glenfield Hospital, Leicester, UK) consists of a free-field microphone and MP3 digital recording device [34] (Figure 4(b)). It also enables 24 h recordings. A keyword spotting method based on hidden Markov models is applied in this system to select possible cough fragments [39]. Then, human experts select some of these cough sounds to develop a statistical model to fit the current record. Finally, the designed model is used to handle the whole recording. As a result, the sensitivity of the system is 91% and the specificity is 99% [40]. The system has been used in clinical trials [41]. Crooks et al. [42] used a hybrid system consisting of the Hull Automatic Cough Counter (HACC) and Leicester Cough Monitor (LCM) software to measure cough frequency during COPD exacerbation convalescence and achieved the overall sensitivity of 57.9% and a specificity of 98.2%.

The VitaloJAKTM system (Vitalograph Ltd., Buckingham, UK, and University Hospital of South Manchester, UK) is a semiautomated cough recording device with two sensors (Figure 4(c)). One is a free-air condenser microphone for manual validation, and the other is a chest wall air-coupled condenser microphone for recording cough sounds [43]. An algorithm based on a median frequency threshold is used to compress 24 h cough sound recordings into average 1.5-hour period. This system is accurate but labor intensive and time consuming because of the manual counting [44]. This system reaches a sensitivity more than 99% in a 24 h ambulatory context on ten patients [43], and it can be used in children to exactly measure cough frequency [45].

Drugman et al. designed an acoustic system using ANNs that was tested on voluntary cough from ten healthy subjects in various circumstances with a sensitivity and specificity of about 95% [46]. In [16], a precise and privacy-protecting cough monitor using a low-cost mobile microphone is proposed by Larson et al. They used principal component analysis (PCA) and a random forest classifier to reconstruct and classify the cough sounds with an average true-positive rate of 92% and a false-positive rate of 0.5% from subjects in the wild. Their system is hence able to protect personal privacy. Amrulloh et al. attempted to design an automated method to automatically identify cough segments from the pediatric sound recordings and achieved a sensitivity and specificity of 93% and 98%, respectively [5].

4.2. Cough Monitors with Mixed Signals. There are two commercial systems using multiple signals to detect cough. The Lifeshirt™ (Vivometrics, San Diego, CA, USA) appeared in 2005. It included several sensors integrated in a shirt worn by the user: electrocardiogram, induction plethysmography, 3-axis accelerometer, and a contact microphone placed on the throat (Figure 4(d)). The device achieved a sensitivity of 78.1% and a specificity of 99.6%. Unluckily, the Lifeshirt is no longer available due to the liquidation of the company in 2009.

The Pulmo Track-CC [47] (KarmelSonix Ltd., Haifa, Israel) includes a piezoelectric belt for monitoring chest wall motion, one lapel microphone, and two contact microphones placed on the trachea and the thorax (Figure 4(e)). The device has been tested over about 2 h in healthy volunteers simulating coughing in different situations (while walking, climbing stairs, and sitting and while in a supine position and in noisy environments). The device achieved a sensitivity of 91% for detecting explosive cough sounds and a specificity of 94% on voluntary cough [47]. However, in a study led by Turner and Bothamley, the device only had a sensitivity of 26% for detecting coughs identified by the ear [35] and it performed very well when detecting coughing caused by acute asthma [48].

4.3. The Ideal Cough Frequency Monitors. The ideal ambulatory cough monitoring system would have these characteristics [3, 49]: mobility, unobtrusiveness, compactness, and privacy protection. More importantly, it can allow 24 hour reliable recording and distinguish cough from other sounds automatically with high sensitivity, high specificity, and proportionately few false-positive events compared to the true-positive events. Audio cough monitoring systems mentioned above have met some of these requirements, but the huge number of noncough sounds limits the development of a cough monitoring system.

5. Cough Sound Processing Algorithms

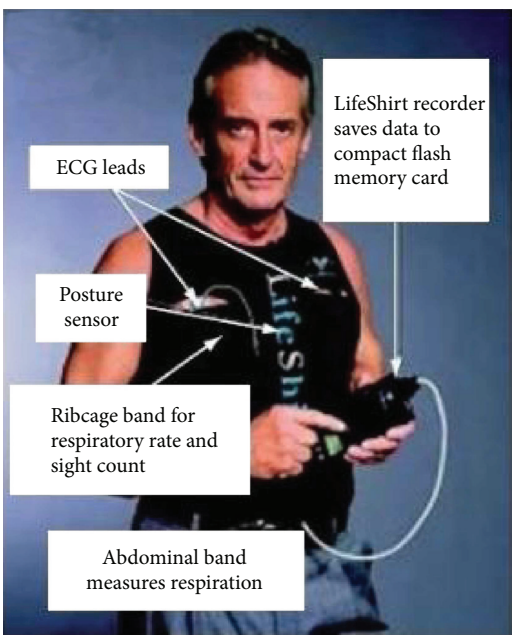
Automatically detecting cough events requires some great answers to at least four major questions [39, 50, 51]: (1) ambient noise reduction: this is an important problem in audio-based detection systems; (2) differentiation from patient sounds, especially speech, laughing, and sneezing: even the most severe cough is far exceeded by the amount of talking; (3) the variability of the cough acoustics: both within and between individuals, combined with the additional complexity of different respiratory diseases; and (4) classification of dry or wet cough: this is a significant medical indicator. Currently, multidisciplinary teams of researchers all over the world are attempting to use pattern recognition techniques such as neural networks, support vector machine (SVM), and naive Bayesian classifier (Bayesian) to manage these questions.

The general workflow for the automatic assessment of cough used in literatures [52, 53] is displayed in Figure 5. The sound signals usually are captured by a microphone, and the first step aims at removing silence within signals. Next, extracting a wide variety of features is a need but this

(a)

(b)

(c)



(d)

(e)

Figure 4: Objective cough monitoring systems: (a) the Hull Automatic Cough Counter; (b) the Leicester Cough Monitor; (c) the VitaloJAKTM system; (d) the Lifeshirt; (e) the Pulmo Track-CC.

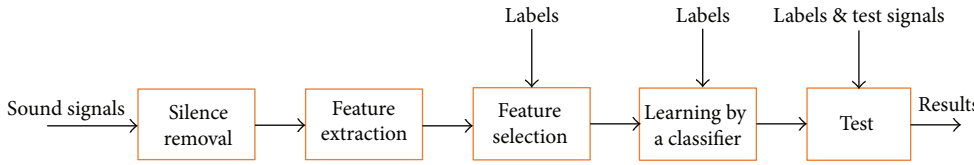


Figure 5: Workflow for the automatic detection of cough.

results in huge amount of data. Therefore, dimensionality reduction is carried out by selecting only the most relevant ones. Finally, a part of the dataset is chosen as training data and they are trained by classifiers. After that, the rest of the dataset is then used for testing.

6. Silence Removal

The raw data of cough sounds contains a lot of silent fragments, whose intensity is low. Removing silence is required to save storage space. In many literatures [25, 37, 54], frame processing is the first and then the start time and end time of cough events are calibrated by the double threshold method using the time-domain features, such as zero-crossing rate and energy entropy.

The energy entropy of a divided audio frame expresses its intensity. E_i can be calculated by following formula:

$$E_i = \sum x_i^2(m), \tag{1}$$

where $x_i(m)$ represents cough sound single after frame processing.

The zero-crossing rate (ZCR) [55] is the ratio of sign changes of a signal. It can enhance the accuracy of the detection of cough sound endpoints. It is defined as

$$\mathrm{ZCR}_i = \sum_m |\mathrm{sign}[x_i(m)] - \mathrm{sign}[x_i(m-1)]|, \tag{2}$$

where $x_i(m)$ represents cough sound single after frame processing and sign $[A]$ is 1 if A is greater than zero and 0 if otherwise.

6.1. Feature Extraction. After silence removal, cough recognition mainly involves extracting features from cough data and then inputting them into a model classifier. Features can be detected from time-domain signals as mentioned above [56] or from frequency-domain signals. Several features have been successfully applied to monitor cough events, including mel-frequency cepstral coefficients (MFCCs) and the characteristic parameters learned by the convolutional neural network (CNN).

6.1.1. Mel-Frequency Cepstral Coefficients (MFCCs). MFCC is based on the hearing mechanism of human beings. The frequency of subjective perception is not linear, and it follows the empirical formula [57]:

$$F_{\mathrm{mel}} = 1125 \log\left(1 + \frac{f}{700}\right), \tag{3}$$

where F_{mel} is the perceptual frequency in mel and f is the actual frequency in hertz.

Therefore, the frequency of cough signals is usually transformed into the perceptual frequency, which can simulate auditory processing better. The concrete steps are as follows [58–61]:

(i) Preemphasize high frequencies, frame, and add windows.

(ii) Take the Fourier transform of each frame signal.

(iii) Calculate the spectral line energy for each frame of data.

(iv) Calculate the logs of the energy at each of the mel frequencies.

(v) Carry out the discrete cosine transform of the results achieved in the fourth step.

(vi) The MFCCs are the coefficient of the results, and usually the first 12 coefficients are taken.

6.1.2. Convolutional Neural Network (CNN). CNN is an efficient identification method which has been developed recently and has attracted extensive attention. Generally, the basic structure of CNN consists of two layers; one is the feature extraction layer. The input of each neuron is connected with the local accepted domain of the previous layer, and the feature of the part is extracted. Once the local feature is extracted and its position relationship between other characteristics can be determined, the other is the feature mapping network layer. Each computing layer is composed of multiple feature maps, each of which is a flat plane, and all neuron weights are equal.

6.2. Learning by Classifiers. There are several classifier algorithms for detecting cough, such as support vector machine (SVM), naive Bayesian classifier (Bayesian), neural network (NN), hidden Markov model (HMM), and dynamic time warping (DTW) [62]. Some studies have been conducted to compare the classifier algorithms with each other [57]. The performance measures are explained as follows [63]:

(1) Accuracy is the percentage of samples correctly classified from the testing data set.

(2) Sensitivity measures the ratio of positives which are exactly identified as well. Sensitivity is defined as the ratio of correctly classified positive samples and true-positive samples.

(3) Specificity measures the percentage of negatives which are recognized intrinsically. Specificity can be calculated as the ratio of correctly classified negative samples and true-negative samples.

The advantages and disadvantages of these algorithms cannot be determined because the results of different experiments are different [34, 64]. However, MFCCs + SVM is used more widely [65, 66], and the neural network has potential to model and achieve accurate identification [67, 68].

7. Future Application of Audio-Based Cough Monitoring

Audio-based cough monitoring has potentially wide application in home medical equipment, clinical trials, and the development of new cough therapies.

7.1. Application in Home Care. Chronic cough is common in old people, and the objective monitoring of chronic cough in the daily life helps to improve the quality of life of the aged with chronic cough [69, 70]. Many doctors stress the importance of early diagnosis of childhood asthma and infantile pneumonia [71–73]. The objective cough assessment provides a probability for this situation. It has been reported that objective cough monitoring is helpful in the diagnosis and treatment of infantile pneumonia [71]. The Pulmo Track-CC produced by KarmelSonix Ltd. achieves a great praise in the diagnosis of asthma [72]. With the development of the cough monitoring device, it could be widely used in home care in the future.

7.2. Response to Treatments and New Antitussives. Clinical treatment trials are a critical part of the diagnosis and management of chronic cough [74]. Some studies have examined the least important difference of cough frequency monitoring that has been available now [24, 75]. Cough monitors are the first choice for the objective evaluation of cough so that they are more widely used in clinical trials as main endpoints. Cough monitors will be a key part of understanding the response of patients with common respiratory diseases.

In recent years, novel antitussives are under development, but the primary outcome measure of antitussive drugs is still subjective, which harms the interests of patients. Many medical literatures point out that a randomized, placebo-controlled, double-blind clinical trial is the gold standard [76, 77]. The primary endpoint should be objective [78, 79]. Objective cough monitoring would be an ideal tool if it can successfully prove the clinical efficacy of novel antitussives. Moreover, subjective outcome measures would be used to assess symptoms and health-related quality of life.

8. Conclusions

Cough is one of the most important symptoms in respiratory clinical trials, while the objective indicators of cough severity are severely absent. This is because the cough frequency, cough intensity, and other objective cough assessment indicators cannot be accurately measured due to technical conditions. This situation has been improved with the development of sound recording and monitoring techniques over the last 20 years.

The generation of cough is not only related to the vocal cords but also to the lungs, and the cough sounds contain a wealth of individual information. Audio-based cough monitors are emerging. In this paper, the basic principle, hardware composition, and experimental results of a cough monitoring instrument are analyzed in detail. This paper also analyzes objective assessment algorithms of cough and their advantages and disadvantages.

Audio-based cough detection systems are now increasingly applied in clinical research. They are becoming more important to study cough. Automated cough algorithms are being developed in quality and processing speed so that audio-based cough monitors will change the assessment of patients' responses to treatments and enter many families in the near future.

Conflicts of Interest

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

References

[1] S. Schappert and C. Burt, "Ambulatory care visits to physician offices, hospital outpatient departments, and emergency departments: United States, 2001–02," *Vital and Health Statistics Series 13, Data from the National Health Survey*, no. 159, pp. 1–66, 2006.

[2] J. Mei, L. Y. Liao, L. Wei, Y. Q. Xie, and K. F. Lai, "Quality assessment of clinical practice guidelines for diagnosis and management of cough in China," *Chinese Journal of Evidence-Based Medicine*, vol. 15, no. 4, pp. 409–413, 2015.

[3] A. H. Morice, G. A. Fontana, M. G. Belvisi et al., "ERS guidelines on the assessment of cough," *European Respiratory Journal*, vol. 29, no. 6, pp. 1256–1276, 2007.

[4] Committee for the Japanese Respiratory Society Guidelines for Management of Cough, S. Kohno, T. Ishida et al., "The Japanese Respiratory Society guidelines for management of cough," *Respirology*, vol. 11, Supplement 4, p. S135, 2006.

[5] Y. A. Amrulloh, U. R. Abeyratne, V. Swarnkar, R. Triasih, and A. Setyati, "Automatic cough segmentation from non-contact sound recordings in pediatric wards," *Biomedical Signal Processing and Control*, vol. 21, pp. 126–136, 2015.

[6] J. Smith and A. Woodcock, "New developments in the objective assessment of cough," *Lung*, vol. 186, no. S1, pp. 48–54, 2008.

[7] S. S. Birring, J. Kavanagh, K. Lai, and A. Chang, "Adult and paediatric cough guidelines: ready for an overhaul?," *Pulmonary Pharmacology & Therapeutics*, vol. 35, pp. 137–144, 2015.

[8] A. Spinou and S. S. Birring, "An update on measurement and monitoring of cough: what are the important study endpoints?," *Journal of Thoracic Disease*, vol. 6, Supplement 7, pp. S728–S734, 2014.

[9] S. S. Birring and A. Spinou, "How best to measure cough clinically," *Current Opinion in Pharmacology*, vol. 22, pp. 37–40, 2015.

[10] P. A. Marsden, J. A. Smith, A. A. Kelsall et al., "A comparison of objective and subjective measures of cough in asthma,"

Journal of Allergy and Clinical Immunology, vol. 122, no. 5, pp. 903–907, 2008.

[11] A. Kelsall, L. A. Houghton, H. Jones, S. Decalmer, K. Mcguinness, and J. A. Smith, "A novel approach to studying the relationship between subjective and objective measures of cough," *Chest*, vol. 139, no. 3, pp. 569–575, 2011.

[12] H. Laciuga, A. E. Brandimore, M. S. Troche, and K. W. Hegland, "Analysis of clinicians' perceptual cough evaluation," *Dysphagia*, vol. 31, no. 4, pp. 521–530, 2016.

[13] H. A. Bickerman and S. E. Itkin, "The effect of a new bronchodilator aerosol on the air flow dynamics of the maximal voluntary cough of patients with bronchial asthma and pulmonary emphysema," *Journal of Chronic Diseases*, vol. 8, no. 5, pp. 629–636, 1958.

[14] W. T. Goldsmith, A. M. Mahmoud, J. S. Reynolds et al., "A system for recording high fidelity cough sound and airflow characteristics," *Annals of Biomedical Engineering*, vol. 38, no. 2, pp. 469–477, 2010.

[15] S. Ren, Y. Shi, M. Cai, and W. Xu, "Influence of secretion on airflow dynamics of mechanical ventilated respiratory system," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. PP, no. 99, p. 1, 2017.

[16] S. Ren, M. Cai, Y. Shi, W. Xu, and X. D. Zhang, "Influence of bronchial diameter change on the airflow dynamics based on a pressure-controlled ventilation system," *International Journal for Numerical Methods in Biomedical Engineering*, 2017.

[17] Y. Hu, E. G. Kim, G. Cao, S. Liu, and Y. Xu, "Physiological acoustic sensing based on accelerometers: a survey for mobile healthcare," *Annals of Biomedical Engineering*, vol. 42, no. 11, pp. 2264–2277, 2014.

[18] E. C. Larson, T. J. Lee, S. Liu, M. Rosenfeld, and S. N. Patel, "Accurate and privacy preserving cough sensing using a low-cost microphone," in *Proceedings of the 13th International Conference on Ubiquitous Computing - UbiComp '11*, Beijing, China, September, 2011.

[19] J. Korpáš, J. Sadloňová, and M. Vrabec, "Analysis of the cough sound: an overview," *Pulmonary Pharmacology*, vol. 9, no. 5-6, pp. 261–268, 1996.

[20] P. Piirilä and A. R. A. Sovijärvi, "Objective assessment of cough," *European Respiratory Journal*, vol. 8, no. 11, pp. 1949–1956, 1995.

[21] J. Korpas, J. Sadlonova, D. Salat, and E. Masarova, "The origin of cough sounds," *Bulletin Europeen de Physiopathologie Respiratoire*, vol. 23, pp. 47s–50s, 1986.

[22] Y. Shi, B. Zhang, M. Cai, and X. D. Zhang, "Numerical simulation of volume-controlled mechanical ventilated respiratory system with 2 different lungs," *International Journal for Numerical Methods in Biomedical Engineering*, vol. 33, no. 9, 2016.

[23] J. Smith, "Ambulatory methods for recording cough," *Pulmonary Pharmacology & Therapeutics*, vol. 20, no. 4, pp. 313–318, 2007.

[24] K. K. Lee, S. Matos, D. H. Evans, P. White, I. D. Pavord, and S. S. Birring, "A longitudinal assessment of acute cough," *American Journal of Respiratory and Critical Care Medicine*, vol. 187, no. 9, pp. 991–997, 2013.

[25] K. K. Lee, A. Savani, S. Matos, D. H. Evans, I. D. Pavord, and S. S. Birring, "Four-hour cough frequency monitoring in chronic cough," *Chest*, vol. 142, no. 5, pp. 1237–1243, 2012.

[26] S. S. Birring, S. Matos, R. B. Patel, B. Prudon, D. H. Evans, and I. D. Pavord, "Cough frequency, cough sensitivity and health status in patients with chronic cough," *Respiratory Medicine*, vol. 100, no. 6, pp. 1105–1109, 2006.

[27] S. S. Birring, "Controversies in the evaluation and management of chronic cough," *American Journal of Respiratory and Critical Care Medicine*, vol. 183, no. 6, pp. 708–715, 2011.

[28] M. Vernon, N. K. Leidy, A. Nacson, and L. Nelsen, "Measuring cough severity: perspectives from the literature and from patients with chronic cough," *Cough*, vol. 5, no. 1, p. 5, 2009.

[29] G. J. Gibson, W. Whitelaw, N. Siafakas, G. S. Supinski, J. W. Fitting, F. Bellemare et al., "ATS/ERS statement on respiratory muscle testing," *American Journal of Respiratory and Critical Care Medicine*, vol. 166, no. 4, pp. 518–624, 2002.

[30] Y. Shi, B. Zhang, M. Cai, and W. Xu, "Coupling effect of double lungs on a VCV ventilator with automatic secretion clearance function," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. PP, no. 99, p. 1, 2017.

[31] L. Pavesi, S. Subburaj, and K. Portershaw, "Application and validation of a computerized cough acquisition system for objective monitoring of acute cough: a meta-analysis," *Chest*, vol. 120, no. 4, pp. 1121–1128, 2001.

[32] K. K. Lee, S. Matos, K. Ward, E. Raywood, D. H. Evans, J. Moxham et al., "P158 cough sound intensity: the development of a novel measure of cough severity," *Thorax*, vol. 67, Supplement 2, pp. A130.3–A1A131, 2012.

[33] A. A. Abaza, J. B. Day, J. S. Reynolds et al., "Classification of voluntary cough sound and airflow patterns for detecting abnormal pulmonary function," *Cough*, vol. 5, no. 1, pp. 8–12, 2009.

[34] S. T. Kulnik, N. M. Williams, L. Kalra, J. Moxham, and S. S. Birring, "Cough frequency monitors: can they discriminate patient from environmental coughs?," *Journal of Thoracic Disease*, vol. 8, no. 11, pp. 3152–3159, 2016.

[35] R. D. Turner and G. H. Bothamley, "How to count coughs? Counting by ear, the effect of visual data and the evaluation of an automated cough monitor," *Respiratory Medicine*, vol. 108, no. 12, pp. 1808–1815, 2014.

[36] J. Smith, "Monitoring chronic cough: current and future techniques," *Expert Review of Respiratory Medicine*, vol. 4, no. 5, pp. 673–683, 2010.

[37] T. Drugman, J. Urbain, N. Bauwens et al., "Objective study of sensor relevance for automatic cough detection," *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 3, pp. 699–707, 2013.

[38] S. J. Barry, A. D. Dane, A. H. Morice, and A. D. Walmsley, "The automatic recognition and counting of cough," *Cough*, vol. 2, no. 1, pp. 8-9, 2006.

[39] J. Amoh and K. Odame, "Technologies for developing ambulatory cough monitoring devices," *Critical Reviews in Biomedical Engineering*, vol. 41, no. 6, pp. 457–468, 2013.

[40] S. S. Birring, T. Fleming, S. Matos, A. A. Raj, D. H. Evans, and I. D. Pavord, "The Leicester Cough Monitor: preliminary validation of an automated cough detection system in chronic cough," *European Respiratory Journal*, vol. 31, no. 5, pp. 1013–1018, 2008.

[41] N. M. Ryan, S. S. Birring, and P. G. Gibson, "Gabapentin for refractory chronic cough: a randomised, double-blind, placebo-controlled trial," *Lancet*, vol. 380, no. 9853, pp. 1583–1589, 2012.

[42] M. G. Crooks, Y. Hayman, A. Innes, J. Williamson, C. E. Wright, and A. H. Morice, "Objective measurement of cough frequency during COPD exacerbation convalescence," *Lung*, vol. 194, no. 1, pp. 117–120, 2016.

[43] K. Mcguinness, K. Holt, R. Dockry, and J. Smith, "P159 validation of the VitaloJAK™ 24 hour ambulatory cough monitor," *Thorax*, vol. 67, Supplement 2, pp. A131.1–A1A131, 2012.

[44] A. Barton, P. Gaydecki, K. Holt, and J. A. Smith, "Data reduction for cough studies using distribution of audio frequency content," *Cough*, vol. 8, no. 1, p. 12, 2012.

[45] D. D. Elghamoudi, H. Sumner, K. Mcguiness, J. Smith, and C. S. Murray, "P241 the feasibility and validity of objective cough monitoring in children using an adult cough detection system," *Thorax*, vol. 70, Supplement 3, pp. A198.1–A1A198, 2015.

[46] T. Drugman, J. Urbain, N. Bauwens et al., "Audio and contact microphones for cough detection," in *Thirteenth Annual Conference of the International Speech Communication Association*, Portland, OR, USA, September, 2012.

[47] E. Vizel, M. Yigla, Y. Goryachev et al., "Validation of an ambulatory cough detection and counting application using voluntary cough under different conditions," *Cough*, vol. 6, no. 1, p. 3, 2010.

[48] R. D. Turner and G. H. Bothamley, "Automated measurement of cough frequency: agreement between auditory analysis and the PulmoTrack" cough monitor in respiratory diseases," in *A37 Cough: Sensitivity, Measurement and Treatment*, p. A1354-A, American Thoracic Society, Philadelphia, PA, USA, 2013.

[49] D. Shen, Z. Qian, and S. Yan, "Dynamic characteristics of mechanical ventilation system of double lungs with bi-level positive airway pressure model," *Computational & Mathematical Methods in Medicine*, vol. 2016, Article ID 9234537, 13 pages, 2016.

[50] C. Pham, "MobiCough: real-time cough detection and monitoring using low-cost mobile devices," in *Asian Conference on Intelligent Information and Database Systems*, Da Nang, Vietnam, March, 2016.

[51] B. H. Tracey, G. Comina, S. Larson, and M. Bravard, "Cough detection algorithm for monitoring patient recovery from pulmonary tuberculosis," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Boston, MA, USA, September 2011.

[52] S. A. Mahmoudi, "Sensor-based system for automatic cough detection and classification," in *ICT Innovations Conference 2016 Element 2015 - Enhanced Living Environments*, Macedonia, October, 2015.

[53] X. U. Weiqing, S. Yan, J. Niu, and M. A. Cai, "Study on air flow dynamic characteristic of mechanical ventilation of a lung simulator," *Science China Technological Sciences*, vol. 60, no. 2, pp. 243–250, 2017.

[54] M. You, H. Wang, Z. Liu et al., "Novel feature extraction method for cough detection using NMF," *IET Signal Processing*, vol. 11, no. 5, pp. 515–520, 2017.

[55] C. H. Chen and D. Stork, "Handbook of pattern recognition & computer vision," *International Journal of Neural Systems*, vol. 5, no. 3, p. 257, 1994.

[56] S. A. Walke and V. R. Thool, "Differentiating nature of cough sounds in time domain analysis," in *2015 International Conference on Industrial Instrumentation and Control (ICIC)*, Pune, India, May, 2015.

[57] S. G. Koolagudi, D. Rastogi, and K. S. Rao, "Identification of language using mel-frequency cepstral coefficients (MFCC)," *Procedia Engineering*, vol. 38, pp. 3391–3398, 2012.

[58] L. Muda, M. Begam, and I. Elamvazuthi, "Voice recognition algorithms using Mel frequency cepstral coefficient (MFCC) and dynamic time warping (DTW) techniques," *Journal of Computing*, vol. 2, no. 3, 2010.

[59] P. M. Chauhan and N. P. Desai, "Mel frequency cepstral coefficients (MFCC) based speaker identification in noisy environment using wiener filter," in *2014 International Conference on Green Computing Communication and Electrical Engineering (ICGCCEE)*, Coimbatore, India, March, 2014.

[60] A. C. Kelly and C. Gobl, "A comparison of mel-frequency cepstral coefficient (MFCC) calculation techniques," *Journal of Computing*, vol. 3, no. 10, pp. 2151–9617, 2011.

[61] C. A. Nurahmad and M. Adriani, "Identifying traditional music instruments on polyphonic Indonesian folksong using mel-frequency cepstral coefficients (MFCC)," in *Proceedings of the 10th International Conference on Advances in Mobile Computing & Multimedia - MoMM '12*, Bali, Indonesia, December, 2012.

[62] B. Ferdousi, S. M. F. Ahsanullah, K. Abdullah-Al-Mamun, and M. N. Huda, "Cough detection using speech analysis," in *2015 18th International Conference on Computer and Information Technology (ICCIT)*, Dhaka, Bangladesh, December 2015.

[63] Y. Liu, Z. Li, and M. Du, "A cough sound recognition algorithm based on time-frequency energy distribution," *Journal of Biomedical Engineering*, vol. 26, no. 5, pp. 953–958, 2009.

[64] A. L. Key, K. Holt, A. Hamilton, J. A. Smith, and J. E. Earis, "Objective cough frequency in idiopathic pulmonary fibrosis," *Cough*, vol. 6, no. 1, pp. 4–7, 2010.

[65] J. Amoh and K. Odame, "DeepCough: a deep convolutional neural network in a wearable cough detection system," in *2015 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, Atlanta, GA, USA, October 2015.

[66] N. Yousaf, W. Monteiro, S. Matos, S. S. Birring, and I. D. Pavord, "Cough frequency in health and disease," *European Respiratory Journal*, vol. 41, no. 1, pp. 241–243, 2012.

[67] V. Swarnkar, U. R. Abeyratne, Y. Amrulloh, C. Hukins, R. Triasih, and A. Setyati, "Neural network based algorithm for automatic identification of cough sounds," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Osaka, Japan, July, 2013.

[68] A. Spinou, R. Garrod, K. Lee et al., "P8 objective cough frequency monitoring in bronchiectasis," *Thorax*, vol. 69, Supplement 2, pp. A80-A81, 2014.

[69] W. Ma, L. Yu, Y. Wang, X. Li, H. Lü, and Z. Qiu, "Changes in health-related quality of life and clinical implications in Chinese patients with chronic cough," *Cough*, vol. 5, no. 1, pp. 1–7, 2009.

[70] W. J. Song, M. Y. Kim, S. E. Lee et al., "E2-6 chronic cough in the elderly population: its impact on quality of life and the association with upper airway symptoms (English session 2)," *Japanese Journal of Allergology*, vol. 61, no. 9-10, p. 1430, 2017.

[71] U. R. Abeyratne, V. Swarnkar, A. Setyati, and R. Triasih, "Cough sound analysis can rapidly diagnose childhood pneumonia," *Annals of Biomedical Engineering*, vol. 41, no. 11, pp. 2448–2462, 2013.

[72] A. L. Boner, G. L. Piacentini, D. G. Peroni et al., "Children with nocturnal asthma wheeze intermittently during sleep," *Journal of Asthma*, vol. 47, no. 3, pp. 290–294, 2010.

[73] K. Hirai, H. Tabata, M. Hirayama, T. Kobayashi, Y. Oh, and H. Mochizuki, "A new method for objectively evaluating childhood nocturnal cough," *Pediatric Pulmonology*, vol. 50, no. 5, pp. 460–468, 2015.

[74] A. H. Morice, L. Mcgarvey, and I. Pavord, "Recommendations for the management of cough in adults," *Thorax*, vol. 61, Supplement 1, pp. i1–i24, 2006.

[75] K. M. Schmit, R. R. Coeytaux, A. P. Goode et al., "Evaluating cough assessment tools: a systematic review," *Chest*, vol. 144, no. 6, pp. 1819–1826, 2013.

[76] A. Morice and P. Kardos, "Comprehensive evidence-based review on European antitussives," *Bmj Open Respiratory Research*, vol. 3, no. 1, article e000137, 2016.

[77] S. Birring, "Developing antitussives: the ideal clinical trial," *Pulmonary Pharmacology & Therapeutics*, vol. 22, no. 2, pp. 155–158, 2009.

[78] Y. Shi, Y. Wang, M. Cai, B. Zhang, and J. Zhu, "An aviation oxygen supply system based on a mechanical ventilation model," *Chinese Journal of Aeronautics*, 2017.

[79] J. Niu, Y. Shi, M. Cai et al., "Detection of sputum by interpreting the time-frequency distribution of respiratory sound signal using image processing techniques," *Bioinformatics*, article btx652, 2017.

Hindawi
Journal of Sensors
Volume 2017, Article ID 7471217, 6 pages
https://doi.org/10.1155/2017/7471217

Research Article

Analysis of Shooting Consistency in Archers: A Dynamic Time Warping Algorithm-Based Approach

Cheng-Hao Quan,[1] Zia Mohy-Ud-Din,[2] and Sangmin Lee[1,3]

[1]*Institute for Information and Electronic Research (IIER), Inha University, Incheon 22212, Republic of Korea*
[2]*Department of Biomedical Engineering, Sir Syed University of Engineering and Technology, Karachi, Pakistan*
[3]*Department of Electronic Engineering, Inha University, Incheon 22212, Republic of Korea*

Correspondence should be addressed to Sangmin Lee; sanglee@inha.ac.kr

Received 25 October 2017; Accepted 2 December 2017; Published 31 December 2017

Academic Editor: Ka L. Man

The shooting consistency of an archer is commonly perceived to be an important determinant of successful scores. Four ($n = 4$) elementary level archers from a middle school in Korea participated in this study. In order to quantify shooting consistency, movement of the bow forearm was measured with an inertia sensor during archery shooting. The shooting consistency was calculated and defined by the dynamic time warping (DTW) algorithm as the distance between two time sequences of acceleration data. Small distance values indicate that the archer has maintained high-level shooting consistency while archery shooting repetitively. To verify the shooting consistency metric, the relationship between scores and shooting consistency is evaluated. The results show that the higher the scores achieved by the archer, the higher is the level of shooting consistency demonstrated.

1. Introduction

Recent advances in microelectromechanical system (MEMS) technologies and wireless communication have allowed the development of small low-cost inertia sensors that can measure body movements precisely and conveniently. The use of inertia sensors to measure activity levels for sporting is emerging as a popular method for the biomechanical quantification of sporting activities (e.g., archery shooting).

Few studies on postural consistency or movement during archery shooting have been found in the literature. To the best of our knowledge, the first researchers to pay attention to postural consistency were Stuart and Atha. In their study [1], Stuart and Atha used a three-dimensional optoelectronic scanner to measure the positions of the archer's head and elbow and that of the bow at the moment of "loose" (releasing the arrow). Postural consistency is defined as the consistency in the archer's positioning that is within-ends the standard deviation (SD) of the positions for each archer.

More recent studies [2, 3] have reported considerable findings with regard to the archer's movement during arrow release. In [2], by comparing the muscular activation patterns of the forearm that manipulates the bow ("bow-forearm") of elite archers and beginners, the study concluded that the action of not contracting both M. flexor digitorum superficialis (MFDS) and M. extensor digitorum (MED), or of only contracting MED, during archery shooting is an indicator of the archer's performance level. In [3], by examining the finger and bowstring movements during arrow release, the study presumed that maximum lateral bowstring deflection does not adversely affect the archer's performance. However, the two aforementioned studies used EMG data to analyze muscular activation patterns or camera-based motion tracking systems to quantify movement; furthermore, the movement or muscular activation only focused on the release moment.

Motion analysis of the movements performed individually by an archer in training or competition provides information about the correctness of the individual techniques and the effectiveness of the archer's strategies. Currently, there are motion tracking systems that use inertia sensors and do not need cameras or stick markers, and they can

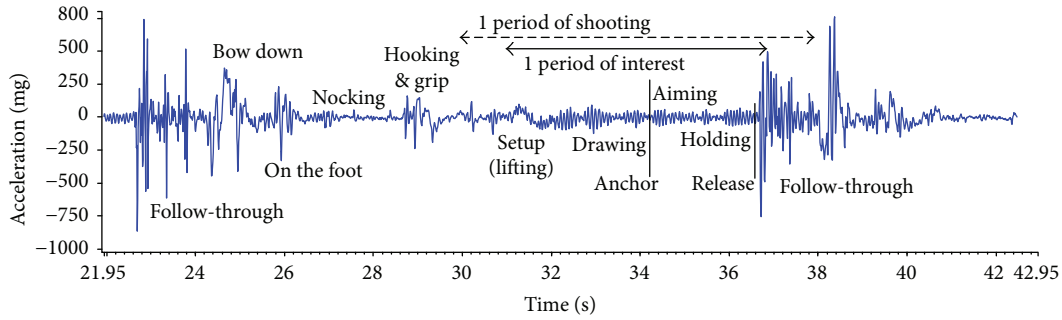


Figure 1: Example of acceleration data (21.95 s–42.95 s): measured bow-forearm movement (x-axis: drawing direction) during successive archery shootings (sampling rate of 100 Hz).

be used anywhere. Such systems also make it possible to analyze body movements during archery shooting more precisely and conveniently than before. The disadvantage is the high cost and need to properly process and interpret measured data.

In this study, we measure bow-forearm movement by using inertia sensors during archery shooting to evaluate the shooting consistency of archers. Moreover, we attempt to provide movement analysis tools that work precisely and conveniently. Shooting consistency is defined as a function of the dynamic time warping (DTW) distance between two time sequences of acceleration data calculated with the DTW algorithm. Small distance values indicate that an archer has maintained high-level shooting consistency while archery shooting repetitively. To the best of our knowledge, no quantifying shooting consistency in archery by using inertia sensors has been reported in the literature. Here, we present the DTW algorithm as a useful signal processing method for quantifying shooting consistency in archery.

2. Materials and Methods

2.1. Archery Shooting. Archery can be described as a comparatively static sport that requires strength and endurance of the upper body, in particular, the forearm and shoulder girdle [4]. Skill in archery is defined as the ability to shoot an arrow onto a given target in a certain time span with accuracy [5]. Archery shooting is described as a three-phase (drawing, aiming, and release) movement. Alternatively, Nishizono et al. [6] further divided the stages of a shot into six: bow hold, drawing, full draw, aiming, release, and follow-through. An archery shot must be well balanced and highly reproducible in order to achieve commendable results in an archery competition.

Figure 1 shows an example of acceleration data that measures the bow-forearm movement during successive archery shootings. As shown in Figure 1, a one-shot period usually consists of nocking, hooking and grip, setup (bow lifting), and drawing that includes anchor moment, aiming, and holding, which in turn includes the release moment, follow through, and finish phases. Not all the phases can be separated exactly, but the phases are meaningful for training.

2.2. Measurement of Acceleration Data. Four elementary level (performance level: beginner, elementary, national, and world-class archer) female participants (middle school student archers) were recorded during archery shooting. The archer's age, previous season's scores, and current scores are as follows: 16, 59, 59.5; 14, 59, 59.5; 15, 55, 53.5; 14, 54, 53.7. The participants were briefed on the purpose of this study and the related procedures.

The participants were required to use their own bow and arrows for measurement purposes and were tested at their school's outdoor tracks. An acceleration data logger (Myomotion, Noraxon Inc., USA) was used to obtain detailed information on body movement over a round (six shots per round required almost 110 s–140 s). An inertia sensor mounted on the bow-forearm was used to measure movement during archery shooting at a sampling frequency of 100 Hz. For each recording round, an archer shot six shots at a distance of 30 m and took a short rest. Each participant shot a total of 36 arrows. Finally, we collected a total of 144 data sets from the four participants ($n = 4$) (considering references [1, 3]).

2.3. Period for Analysis. In this study, for the analysis of shooting consistency in archery, we set the period of interest (POI, see Figure 1) that only includes some interesting phases, as follows: setup (not the whole period), drawing, aiming, holding, and a period of dozens of milliseconds after release.

As MEMS technology has developed, the ability to measure body movements more precisely and conveniently than before has increased. For the analysis of shooting consistency in archery, we choose wider range of POI than previous works. Such that it provides more useful information compared with studies that have focused on release moments only.

2.4. Method for Analysis of Shooting Consistency. In order to analyze shooting consistency, the bow-forearm movement is measured during archery shooting, and the DTW distance between two time sequences of acceleration data is calculated using the DTW algorithm.

DTW is a well-known technique for finding an optimal nonlinear mapping between two given time sequences (e.g., acceleration data in archery shooting) under certain restrictions [7]. The objective of DTW is to compare two time sequences. The distance (DTW distance) between X and Y indicates how much similarity there is between them.

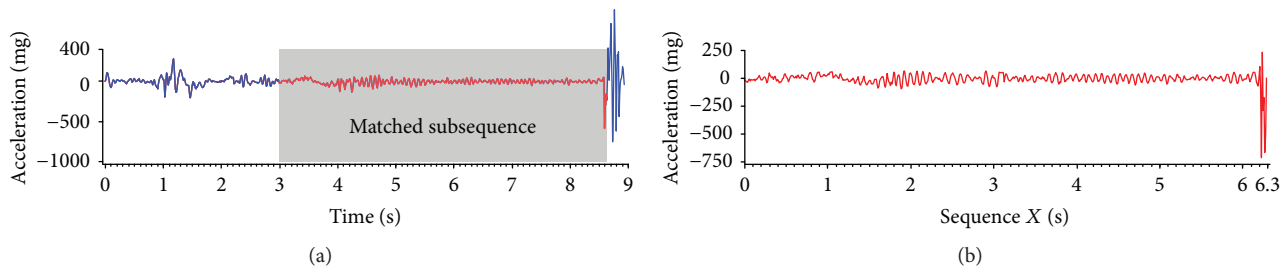


FIGURE 2: (a) Example of input sequences for DTW: query sequence Y (109.21 s long) for one-round acceleration data and reference sequence X (6.3 s long) within POI indicated by grey region. (b) Zoom-out of sequence that corresponds to sequence X of (a).

In fields of speech recognition [8], and data mining and information retrieval [9–11], DTW has been successfully applied to cope with time deformations and different speeds associated with automatic time-dependent data.

In essence, DTW is an algorithm that can compute the similarity between two time sequences, even if the lengths do not match [7, 12]. Assume that we have two time sequences, $X = (x_1, x_2, \dots, x_n)$ of length $n \in N$ and $Y = (y_1, y_2, \dots, y_m)$ of length $m \in N$. In order to measure the similarity between these two time sequences, we construct an n-by-m cost matrix where the $(i^{\text{th}}, j^{\text{th}})$ element of the matrix contains cost $c(x_i, y_j)$ between two points x_i and y_j. Typically, $c(x_i, y_j)$ is small (low cost) if x_i and y_j are similar to each other; otherwise, $c(x_i, y_j)$ is large (high cost). The cost matrix of X and Y using the Manhattan distance (absolute value of the difference) is a local cost measure c (i.e., $c(x_i, y_j) = |x_i - y_j|$). An (n, m)- warping path $p = (p_1, p_2, \dots, p_l)$ of length $l \in N$ defines the mapping between the two time sequences X and Y. The k^{th} element of p is defined as $p_k = (i, j)_k$. There are exponentially many warping paths. However, we are only interested in the path that minimizes the cost (warping cost) with respect to the local cost measure c, which is defined as the sum of all local costs, where k runs from 1 to l. An optimal warping path between X and Y is a warping path p^* with minimal total cost among all possible warping paths. The distance (DTW distance) $d(i, j)$ between X and Y is then defined as the total cost of p^*. This optimal warping path can be found efficiently by using dynamic programming to evaluate the following recursive steps: $d(i, j) = c(i, j) + \min\{d(i-1, j-1), d(i-1, j), d(i, j-1)\}$. The value $d(i, j)$ defines an n-by-m accumulated cost matrix. The initialization can be simplified by extending the accumulated cost matrix with an additional row and column and formally setting $d(i, 0) = \infty$ for $i \in [1 : n]$, $d(0, j) = \infty$ for $j \in [1 : m]$, and $d(0, 0) = 0$. Furthermore, the optimal warping path $p^* = (p_1, p_2, \dots, p_l)$ is computed in reverse order of the indices starting with $p_l = (n, m)_l$.

2.5. DTW with Finding Subsequences Automatically. In our study, DTW is used not only to compare two similar time sequences but also to find subsequences within the longer sequence that optimally fits the shorter sequence. The longer sequence represents a given one-round (six shots) acceleration data that we call "query sequence," and the shorter sequence represents a given acceleration data of POI (see Figure 1) that we call "reference sequence." Before starting DTW, we select one sequence as the reference sequence where the archer achieved a score of 10 for shooting. The problem of finding the optimal subsequences can be solved by a variant of DTW that we call "aDTW" (DTW with finding subsequences automatically).

We illustrate the aDTW algorithm with the example described by figures. The input for aDTW consists of the query (sequence Y in Figure 2(a)) and reference (sequence X in Figures 2(a) and 2(b)).

In order to find a subsequence that minimizes the distance to the reference over all possible subsequences of query, we modify the initial conditions of the classic DTW algorithm by setting $d(i, 0) = \infty$ for $i \in [0 : n]$ and $d(0, j) = 0$ for $j \in [0 : m]$. In the first iteration, we compute the accumulated cost matrix shown in Figure 3(a) and obtain the distance function (Figure 3(b)) that corresponds to the top row of the matrix.

By setting the threshold manually for each participants, we obtain the six local minima with distance below the threshold. The indices of the six local minima correspond to the end point of the six subsequences in query sequence Y. The six blue small circles shown in Figure 3(b) indicate the six local minima. To determine the indices of the starting point of the six subsequences in query sequence Y, the optimal warping path is computed in reverse order of indices starting with each six local minima. The six resulting "matched subsequences," that is, the subsequences of query Y similar to reference X, are shown in Figures 4(a) and 4(b). In the next iteration, using classic DTW, we compute the six final accumulated cost matrices, as shown in Figure 4(a), and finally, normalization of the distance (see Figure 4(b)) between the reference and the matched six subsequences in query is conducted.

3. Results and Discussion

For our time sequence data sets, we only employed the x-axis acceleration data to calculate distance (DTW distance) by using the aDTW algorithm described in Section 2. We iteratively found the 36 matched subsequences for all six successive rounds, but not all participant data showed good matching. The numbers of matched subsequences for the four participants are 35, 35, 29, and 31. One of the 36 subsequences was a reference, and 35 was the maximum number of matched subsequences found automatically.

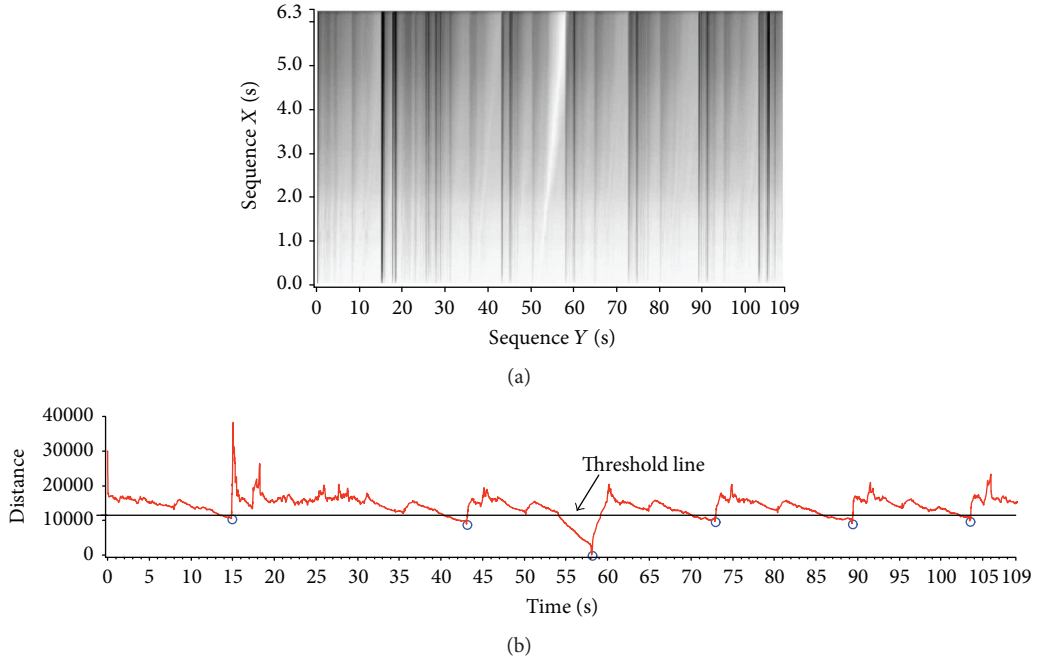


Figure 3: (a) Accumulated cost matrix between sequences X and Y for calculating distance function. Regions of low cost are indicated by light colors, and those of high cost are indicated by dark colors. (b) Distance function that corresponds to top row of accumulated cost matrix. Blue small circles indicate six local minima with distance below threshold. Indices of six local minima correspond to end point of six subsequences in Y.

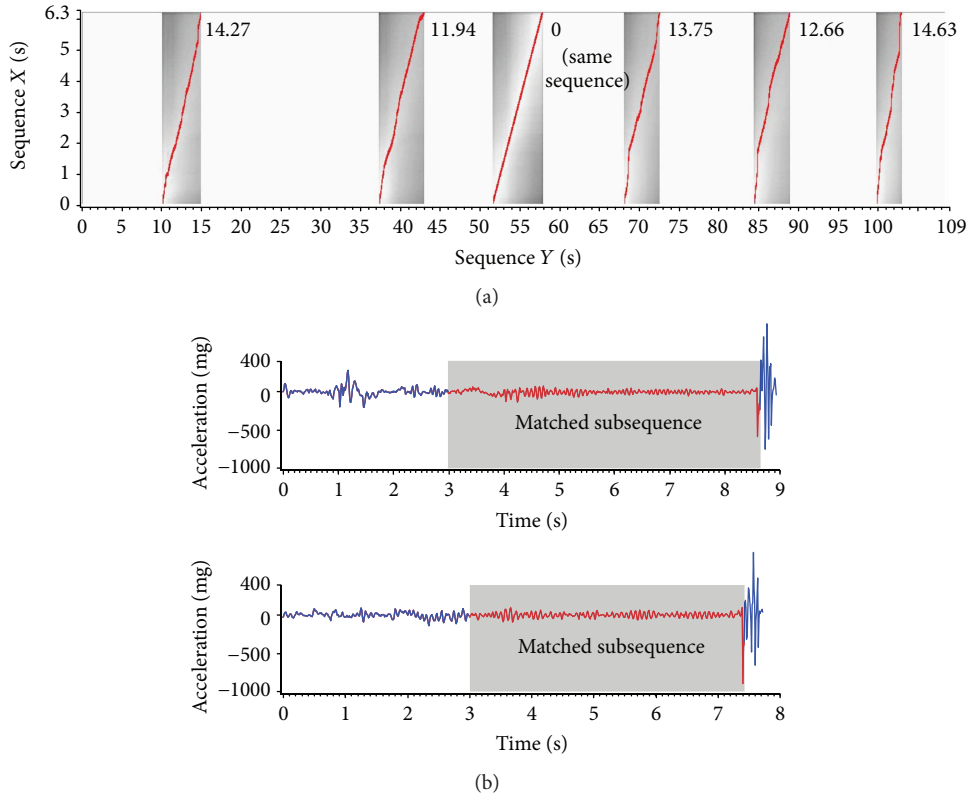


Figure 4: (a) Final accumulated cost matrix between sequence X and six matched subsequences in Y. Regions of low cost are indicated by light colors, and those of high cost are indicated by dark colors. Six grey regions correspond to six matched subsequences in Y. Grey (red in color print) lines indicate six optimal warping paths. Value indicates normalized distance between sequence X and each matched subsequences. (b) Resulting two (correspond to the second and fourth subsequences) of six subsequences indicated by grey (red in color print) lines.

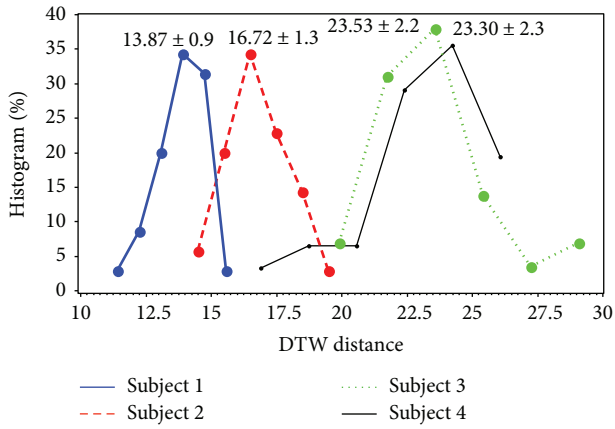


Figure 5: Histogram of distance (DTW distance) that corresponds to four participants and relationship between scores and distance (postural consistency).

Then, we calculated the distance between the reference and the subsequence found automatically for the four participants. The mean and SD value of the distance for each participant are as follows: 13.87 ± 0.9, 16.72 ± 1.3, 23.53 ± 2.2, and 23.30 ± 2.3.

In order to evaluate the relationship between distance and the scores obtained, we plotted a histogram of the DTW distance of each participant. The histogram of DTW distances that correspond to the four participants and the relationship between shooting consistency (distance) and scores are presented in Figure 5. The results show that the higher the score (59 of 60) achieved by an archer, the higher level is the shooting consistency (lower distance, that is, distance of 13.87 and 16.72) that they showed. Subjects 1 and 2 are female archers with relatively higher score. A small distance value indicates that an archer has maintained high-level shooting consistency while archery shooting repetitively. On the other hand, the lower the score (54 of 60 and 55 of 60) achieved by an archer, the lower level is the shooting consistency (higher distance, that is, distance of 23.53 and 23.30) that they showed. Subjects 3 and 4 are female archers with relatively lower score.

In this study, we provide a useful signal processing method that quantifies shooting consistency in archery. The results show that the DTW algorithm is a very useful metric for assessing shooting skills, evaluating the progress on training, and finding talented archers.

4. Conclusions

In this study, we measured four ($n = 4$) elementary level archers from a middle school in Korea for the purpose of analyzing shooting consistency in archery. Such shooting consistency is defined by the distance between two time sequences with respect to acceleration data. The distance was calculated using the DTW algorithm. The results showed that the higher the scores achieved by an archer, the lower is the distance, that is, the higher level is the shooting consistency showed by the archer.

Consequently, distance can be used as a quantitative parameter for measuring similarity of movement, which is, shooting consistency in archery. Moreover, the DTW algorithm-based approach provides the advantage of searching all matched subsequences automatically from a long sequence. We inferred that the proposed approach might be important for assessing shooting skills, evaluating an archer's progress, and finding talented archers in advance. The ultimate goal of this study is to develop measures to identify archers' talents and using these measures to improve archers' performance and preventing injuries ahead.

Conflicts of Interest

The authors declare no conflict of interest.

Authors' Contributions

Cheng-Hao Quan and Sangmin Lee conceived and designed the experiments. Cheng-Hao Quan and Sangmin Lee performed the experiments. Cheng-Hao Quan analyzed the data and prepared the figures. Cheng-Hao Quan drafted the paper. Cheng-Hao Quan, Zia Mohy-Ud-Din, and Sangmin Lee revised and approved the final version of the paper.

Acknowledgments

This research was supported by the Basic Science Research Program through the NRF (National Research Foundation of Korea) funded by the Ministry of Education (2010-0020163).

References

[1] J. Stuart and J. Atha, "Postural consistency in skilled archers," *Journal of Sports Sciences*, vol. 8, no. 3, pp. 223–234, 1990.

[2] H. Ertan, "Muscular activation patterns of the bow arm in recurve archery," *Journal of Science and Medicine in Sport*, vol. 12, no. 3, pp. 357–360, 2009.

[3] B. Horsak and M. Heller, "A three-dimensional analysis of finger and bow string movements during the release in archery," *Journal of Applied Biomechanics*, vol. 27, pp. 151–160, 2011.

[4] D. L. Mann and N. Littke, "Shoulder injuries in archery," *Canadian Journal of Sport Sciences*, vol. 1989, no. 14, pp. 85–92, 1989.

[5] P. Leroyer, J. Van Hoecke, and J. N. Helal, "Biomechanical study of the final push-pull in archer," *Journal of Sports Sciences*, vol. 11, no. 1, pp. 63–69, 1993.

[6] A. Nishizono, H. Shibayama, T. Izuta, and K. Saito, *Analysis of Archery Shooting Techniques by Means of Electromyography*, J. Tsarouchas, J. Terauds, B. A. Gowitzke and L. E. Holt, Eds., International symposium on biomechanics in sports, Athens, Greece, 1987.

[7] E. Keogh and A. Ratanamahatana, *Everything You Know About Dynamic Time Warping Is Wrong*, Workshop on Mining Temporal and Sequential Data, in conjunction with 10th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD-2004), Seattle, WA, USA, 3rd edition, 2004.

[8] F. Itakura, "Minimum prediction residual principle applied to speech recognition," in *IEEE Transactions on Acoustics, Speech, and Signal Processing*, pp. 52–72, IEEE, New York, NY, USA, 1975.

[9] M. H. Ko, G. West, S. Venkatesh, and M. Kumar, "Using dynamic time warping for online temporal fusion in multisensor systems," *Information Fusion*, vol. 9, no. 3, pp. 370–388, 2008.

[10] J. Liu, L. Zhong, J. Wickramasuriya, and V. Vasudevan, "Uwave: accelerometer-based personalized gesture recognition and its applications," *Pervasive and Mobile Computing*, vol. 5, no. 6, pp. 657–675, 2009.

[11] N. Gillian, R. B. Knapp, and S. O'Modhrain, "Recognition of multivariate temporal musical gestures using N-dimensional dynamic time warping," in *The 11th Int'l conference on New Interfaces for Musical Expression*, Oslo, Norway, 2011.

[12] E. Keogh and A. C. Ratanamahatana, "Exact indexing of dynamic time warping," *Knowledge and Information Systems*, vol. 7, pp. 358–386, 2004.

Hindawi
Journal of Sensors
Volume 2017, Article ID 7483075, 9 pages
https://doi.org/10.1155/2017/7483075

Research Article

The Intelligent Healthcare Data Management System Using Nanosensors

Ulzii-Orshikh Dorj,[1] Malrey Lee,[1] Jae-young Choi,[2] Young-Keun Lee,[3] and Gisung Jeong[4]

[1]*Center for Advanced Image and Information Technology, School of Electronics & Information Engineering, Chonbuk National University, 664-14, 1Ga, DeokJin-Dong, Jeonju City, Chonbuk 561-756, Republic of Korea*
[2]*Department of Computer Engineering, Sungkyunkwan University, Seoul City, Republic of Korea*
[3]*Department of Orthopedic Surgery, Research Institute of Clinical Medicine of Chonbuk National University, Biomedical Research Institute of Chonbuk National University Hospital 20, Geonji-ro, Jeonju 54907, Republic of Korea*
[4]*Department of Fire Service Administration, Wonkwang University, Iksan City, Republic of Korea*

Correspondence should be addressed to Malrey Lee; mrlee@chonbuk.ac.kr, Jae-young Choi; jaeychoi@skku.edu, Young-Keun Lee; trueyklee@naver.com, and Gisung Jeong; jgskor@wonkwang.ac.kr

Received 27 July 2017; Accepted 15 October 2017; Published 25 December 2017

Academic Editor: Mucheol Kim

We developed a design of Intelligent Healthcare Data Management System using nanosensors (IHDMS) and composed an application for mobile device. The proposed IHDMS will coordinate the healthcare data of the patients from nanosensors and transforms it into a worldwide consumed standard HL7 (Health Level Seven) for conversion of healthcare data. This converted data dispatches to a server of its system. The battery lifetime of the facility is feasible to increase, the memory usage is less than 100 KB, and it operates all data by employing few and far between resources. Moreover, the proposed system decreases the waiting time in the transposing data, and secured channel was used for the server of the healthcare center in the running HL7 format data.

1. Introduction

As with acceleration in developing technologies, intelligence healthcare systems are also moving towards novel approaches and models of healthcare based on nanosensor, smart phone, smart watch, and so on. Currently, nanotechnology is getting one of the incredible studies and plays an important role in many fields, such as electronics, telecommunication, agriculture and food, materials, energy storage, biotechnology, medicine, and healthcare system.

Intelligent Healthcare Data Management System allows a sick person or patient to use a healthcare service such as a diagnosis service, emergency management service, and monitoring service whenever and anyplace. The Intelligent Healthcare Data Management System (IHDMS) with nanosensors contains three components. Schematic of the Healthcare Data Management System using nanosensors is illustrated in Figure 1.

Component 1 involves wearable nanosensors worn by patient. Each nanosensor can recognize, represent, and process one or more physiological signals, namely, heart rate monitoring by an electrocardiogram (ECG) nanosensor, brain electrical activity observation by an electroencephalogram (EEG) nanosensor, muscle activity monitoring by an electromyogram (EMG) nanosensor, blood pressure recording by a blood pressure nanosensor, trunk position monitoring by a tilt nanosensor, and so on.

Component 2 involves an individual application, which is executing on a smart phone. This application is answerable for a considerable commitment. It greatly performs as an interface to the wireless medical sensors, which assembles the sick person or patient's health information and transfers it to a medical healthcare server. The WBAN (wireless body area network) contains network configuration and management.

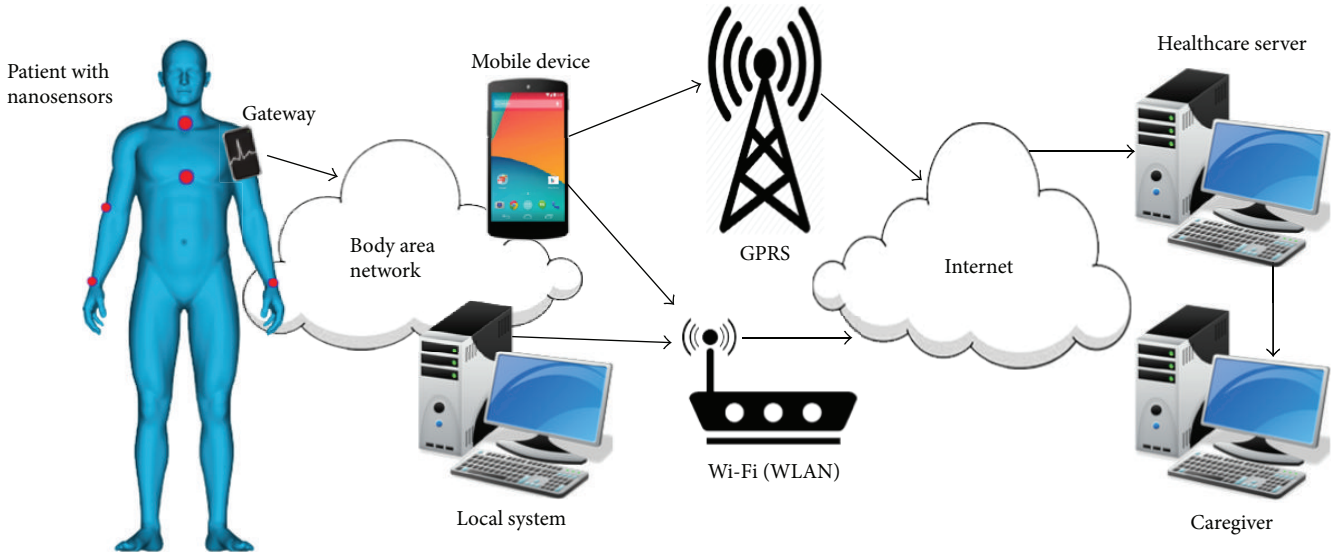


Figure 1: Schematic of the Intelligent Healthcare Data Management System using nanosensors.

The configured WBAN network is conducted by an application. It services to share channels, synchronize time, and also retrieve, process, and transfer of data. Based on the information from multiple medical sensors, the application should determine the patient's health state and provide feedback through a user friendly and intuitive graphical. Consequently, if a communication channel to the medical server is available, the application sets out a secure link to the medical server and sends reports that can be united into the patient's medical record in HL7 format. However, if a link between the mobile device and the medical server is not available, the application stores the data locally and initiates data uploads, whenever a link becomes available.

Component 3 consists a medical server(s) or healthcare server, accessed through the Internet. Moreover, this component could embrace other servers, such as informal caregivers, commercial healthcare providers, and even emergency servers. The medical server commonly composes a service that establishes a communication channel to the patient's application. It assembles the data from the patient application and unites the data into the patient's medical record. The service can allowance admonitions, if reports seem to manifest an anomalous condition.

Numerous researches have been studied in intelligent healthcare system and applications of the nanotechnology. Also, genetic algorithms are already a key player in task assignment. Savić et al. [1] presented a genetic algorithm approach to solve task assignment problem. Lee and Shin [2] developed a task assignment problem in minimizing costs for execution and communication. Mei et al. [3] proposed a task distribution framework to support dynamic reconfiguration of PHMS, by means of task redistribution. This framework consists of a coordinator and a set of facilitators. Task assignment problem in arbitrary processor networks was studied in [4]. Pawar et al. [5] presented a method in optimal task assignment processing based on a genetic algorithm. Bachouch et al. [6] composed an optimization model for task assignment in healthcare.

Hu and Marculescu [7] and Alsalih et al. [8] developed models of the channel directions in processor network. Otto et al. [9] studied about the "Personal Server" in their research work. Paré et al. [10] published a review of home telemonitoring in chronic diseases.

Kang et al. [11] classified stress status based on EEG signals. Sun-Jin [12] designed a smart application that can accurately diagnose and process the current state of the local environment, objects, and persons remotely. Also, Kim et al. [13] developed mobile healthcare application based on the tongue diagnosis.

Kumar et al. [14] proposed a general real-time healthcare monitoring system architecture and summarized the need, challenges, and radio technologies of WBANs, since it is the key factor to achieve low power and low latency. Hanen et al. [15] proposed a mobile medical web service system and implemented a medical cloud multiagent system using Android operating system. Hamdi et al. [16] introduced other system in the management of medical technology. O'Donoghue and Herbert [17] presented the Data Management System (DMS) architecture, which employs an agent-based middleware to intelligently and effectively manage all pervasive medical data sources. Lee et al. [18] composed a management system, which is suitable for diabetes patients by their blood sugar level. Ketan et al. [19] developed a healthcare system for the diabetes patients. Nkenyereye and Jang [20] presented a design in the healthcare web service gateway of the healthcare monitoring system for distance runner.

Sahoo et al. [21] introduced the chief scientific and technical aspects of nanotechnology and discussed some of its potential clinical applications. Omanović-Mikličanin et al. [22] composed the perspective of the nanotechnology applications in medicine. Raffa et al. [23] summarized the progress made in nanotechnology for healthcare. Ali and Abu-Elkheir [24] overviewed a foresight of the healthcare environs and its structural challenges, which are used nanotechnology and nanonetworks.

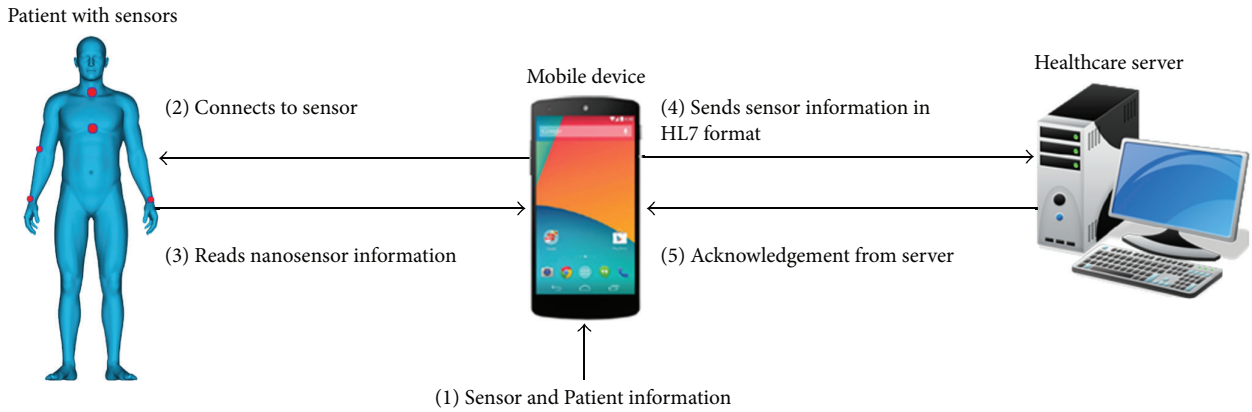


Figure 2: Design of the IHDMS.

Earlier systems were concentrated on obtaining the patient's information irrespective of the resources been used. A raw data produced by these systems were transferred to healthcare server, which caused a lot of overhead of handling this information on the server.

Even though previous researches are composed of various methodologies in the healthcare management system, that is, [14, 17–20], the task assignment, that is, [1–6], the healthcare application for the smart phone, that is, [12, 13, 15], and nanotechnology application for healthcare, that is, [21–24], any work does not perform using these methods together. This study focuses on to developing Intelligent Healthcare Data Management System using nanosensors, which includes task assignment algorithms based on a genetic algorithm, a smart device application, and nanosensors with high sensitivity all together. In addition, genetic algorithm can decrease waiting time in data transpose and demands lower memory while providing the assignment. This paper is arranged as follows. Section 2 explains the objectives and newly proposed methodology. Section 3 presents results and discussion generated from this research work. Section 4 gives conclusions and future works form this study, like the goals achieved and efficiency of the system.

2. Objectives and Methodology

The main goal of this study was to design the Intelligent Healthcare Data Management System using nanosensors and develop an application for mobile device. In our proposed system, a device is ubiquitous, and software supplies the real-time patient health information to the caregiver using nanosensors.

An algorithm to obtain the overall methodology used in this paper is as follows:

(i) Design the Intelligent Healthcare Data Management System (IHDMS) using nanosensors.

(ii) Develop a task assignment algorithm (TAA).

(iii) Transform sensor data to HL7.

(iv) Develop an IHDMS application.

2.1. Task Assignment. In the Healthcare Data Management System, a "single task" is a connection between mobile system and a body sensor at a distinct time and assemble the sick person or patient's physiological health information.

If a TAA commends an individual task to the sensor, operation in that begins. In our composed system, the IHDMS connects to target sensor, assembles, and operates the patient's health information, then keep in present information.

In this research, the TAA was developed using genetic algorithm. The genetic algorithm is found on not complicated evolutionary process and consuming pointed option to get optimal results. In genetic algorithm, population of the chromosome was produced.

2.2. Transform of the Nanosensor Data to HL7. After operating the sensor information, it was transformed to the healthcare server by HL7 standard.

An individual HL7 message consists of MSH, PID, and OBX segments. Therefore, these three kinds of segments are produced in our proposed IHDMS.

2.3. An Application for IHDMS. An application programming interface (API) for smart device was composed in the proposed Intelligent Healthcare Data Management System using nanosensors. The IHDMS application based on Android OS and utilized Java program.

The application consists a main page and other subpages, which are unable to get and enter information about sensors, and patient's health information.

3. Results and Discussion

3.1. Design of the Intelligent Healthcare Data Management System Using Nanosensors. This section presents the design and performance of the Intelligent Healthcare Data Management System (IHDMS) using nanosensors. In this study, a sick person or patient's health information was assembled from the nanosensors. The assembled health information was operated on the phone and transformed to the HL7 format for transfer to the healthcare server. Figure 2 illustrates the design of the IHDMS for better understanding.

3.2. Task Assignment Algorithm (TAA). Task assignment algorithm (TAA) is the key component of this study. Primarily, connection between the nanosensor and the mobile device must be set up. There are several predefined nanosensors, which are the user or the caregiver can choose and connect manually to them. User, who is the first time addressed to connect to the nanosensor, should get permission to make a contact in further by an allowed passport. The allowed passport is saved in the system for user's automatic connection afterwards.

Nanosensor information page for connected nanosensors was displayed. In this page, the user can choose scan time and scan daily inputs. Entered information is saved and required in the TAA to produce chromosomes. In our system, input scan time was determined between 10 to 60 seconds, and scan daily input was limited 24 scans daily. This limitation of the scan time and scan daily is reasoned from the current research demands, such as reduce memory usage and save battery of the mobile device.

If there are no nanosensor inputs (scan time and daily scanned inputs) of the system, any chromosome would not be produced by the TAA in the IHDMS. In this case, default values for every inputs are equal to zero (0). Instantly, the application is operated by the user; the TAA is created.

The chromosomes, which are produced by the TAA, consist two variants of chromosomes with distinct bit sizes. These chromosomes are scanned daily and a single hour. In order to produce the chromosomes, the TAA serves two different functions.

The first function operates once in around the clock and produces the daily scan for each sensor. The second function operates every hour and produces the scan per hour chromosome for the whole activated sensors in the current time.

This scan information can be used to produce the scans per hour chromosome for all nanosensors. Twenty-four bit chromosomes will be produced during this process for each sensor. These produced chromosomes create nanosensor's scan information for every hour of the single day. As a single day divided by twenty-four hours, the size of the scans in each day chromosome is equal to twenty-four bit, and the maximum number of the scans in each day for each nanosensor is also twenty-four. So a cell of the chromosome represents a single hour of the day. The cell of the chromosome can get a value 1 or 0. When the process of the sensor information in the current hour finished, the cell gets a value of 1 bit. When the sensor information will not be processed in the current hour, the cell gets a value of 0 in the chromosome.

Producing these chromosomes is little various from the genetic algorithm-based chromosome generation. In the genetic algorithm, a population of the chromosome was generated, and their fitness was evaluated. Then a novel population through the genetic operations was created. This algorithm is convenient on a powerful system, and there is no limitation on system resources such as memory and battery. The proposed TAA is a genre of the genetic algorithm.

According to the genetic algorithm theory, randomized population of the chromosomes was produced and the chromosomes were rejected based on its fitness. Genetic algorithm for the proposed task assignment is presented as a flowchart in Figure 3.

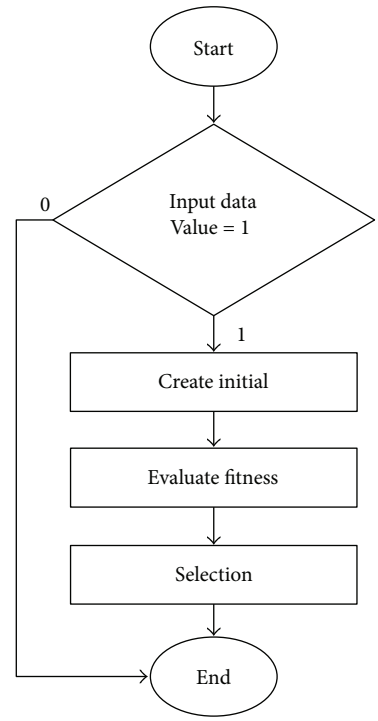


Figure 3: Genetic algorithm for task assignment.

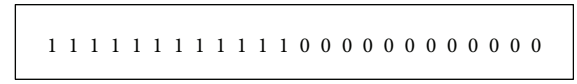


Figure 4: Chromosome before randomization.

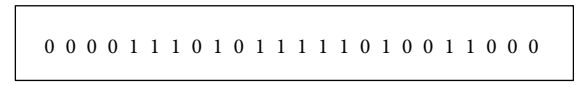


Figure 5: Chromosome after randomization.

This process consists n number of iterations, randomizations, and adoption. In order to reduce the n number of iterations, randomization, and selection, the TAA was started working from the fitness function of the system.

Figure 4 presents produced chromosomes before randomization.

According to the requirement of the system to obtain randomized health information of the patient, the produced chromosomes were randomized in this study. Figure 5 illustrates produced chromosomes after randomization.

The proposed TAA does not have an assumption of crossover and mutation as in the genetic algorithm. The genetic algorithms generally consist of three fundamental operators: reproduction, crossover, and mutation. The crossover and mutation are utilized to produce a new population of the chromosomes, and a fitness function is used to reject unfit chromosomes. The genetic algorithm is utilized to get the best solution after n number of population generation, fitness selections, crossover, and mutation without specifying

BP nanosensor

1	1	0	0	0	1	0	1	0	1	1	1	0	0	0	1	0	1	0	0	0	1	0	0

SPO2 nanosensor

0	1	0	1	0	1	0	0	1	1	1	0	0	0	1	0	1	1	1	0	0	1	0	1

ECG nanosensor

1	1	0	0	0	0	1	1	1	0	0	0	1	0	1	1	0	0	0	1	1	0	1	0

Temperature nanosensor

0	0	0	1	0	0	0	1	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1

Figure 6: Result of scan chromosomes a day in all nanosensors. BP: blood pressure nanosensor; SPO2: saturation of peripheral oxygen nanosensor; ECG: electrocardiogram nanosensor.

any rules. The task assignment algorithm just produced chromosomes. A single chromosome provided by the scans every single day and various chromosomes are produced on all process.

An actual produced chromosome presents the hour information. The chromosomes were produced for each nanosensor based on their input of scan daily. Some produced chromosomes for each nanosensor are shown in Figure 6.

Nanosensor scan information is presented as an array of vectors in Table 1. The array size is 24 (hour) in each vector consisting nanosensors name. This produced array was consumed for the production of second unit of chromosomes.

The second function operates each single hour and produces a chromosome of size 60. It produces a chromosome, which has 60 cell size using the information from array of vectors. Each cell consists a value of 0 or nanosensor name. When the cell value is equal to nanosensor name, the IHDMS operates the nanosensor information.

The produced chromosome at the fifth hour for ECG nanosensor is presented as below.

[ECG Nanosensor, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]

Next step of the production of primary chromosome is to randomize as below.

[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ECG Nanosensor, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]

In the proposed IHDMS, time suspension worked for the chromosomes. Then the system was stopped over until continuation of that process. Due to this stopover in the system, battery life and memory were saved of the device.

The produced chromosomes, which are used for the scan of the nanosensor information, are exclusive, and the nanosensors scan at an exclusive time every single day. It also saves the memory usage and battery and more efficiently operates the nanosensor information on the mobile environment.

3.3. Transform Nanosensor Data to HL7. The nanosensor information was converted into the lab HL7 format. Inbound lab messages can be received as ORU (observational report—unsolicited) messages and will be sent to the "Labs" tab on patient's profile in healthcare records. Each HL7 message consists of one or more segments (see Table 2). A segment is a group of fields, which contain different data types.

MSH field information is illustrated in Table 3: Seq. is a sequence, in which each field information has to be entered and each field is separated by delimiter. Opt. is options, which is optional field by user. Each involved field information is obligated and cannot be left blank. Len. presents the maximum length of the each field.

Table 1: Nanosensor scan information per hour.

Hours	Nanosensors
1	SPO2 nanosensor
2	SPO2 nanosensor, temperature nanosensor
3	Temperature nanosensor
4	BP nanosensor
5	ECG nanosensor
6	SPO2 nanosensor, ECG nanosensor, temperature nanosensor
7	BP nanosensor, ECG nanosensor, temperature nanosensor
8	
9	SPO2 nanosensor, ECG nanosensor
10	
11	SPO2 nanosensor, ECG nanosensor, temperature nanosensor
12	BP nanosensor, ECG nanosensor
13	SPO2 nanosensor, ECG nanosensor, temperature nanosensor
14	ECG nanosensor, temperature nanosensor
15	BP nanosensor, SPO2 nanosensor, ECG nanosensor
16	BP nanosensor, SPO2 nanosensor
17	ECG nanosensor, temperature nanosensor
18	SPO2 nanosensor, temperature nanosensor
19	
20	BP nanosensor, SPO2 nanosensor, ECG nanosensor
21	ECG nanosensor
22	BP nanosensor
23	BP nanosensor, SPO2 nanosensor
24	BP nanosensor, temperature nanosensor

Table 2: HL7 lab messages format.

Segment	Description
MSH	Message header
PID	Patient identification
PV1	Patient visit identification
ORC	Common order
NTE	Notes and comments
OBR	Observation request
OBX	Observation/result
NTE	Notes and comments

Table 3: MSH segment field information.

Seq.	Len.	Fmt	Opt.	Field name
0	3		R	Segment ID = "MSH"
1	1	ST	R	Field separator
2	4	ST	R	Encoding characters
3	20	HD	R	Sending application
4	20	HD	R	Sending facility
5	20	HD	R	Receiving application
6	20	HD	R	Receiving facility
7	14	TS	R	Date/time of message
8	40	ST	O	Security
9	7	CM	R	Message type
10	20	ST	R	Message control ID
11	3	PT	R	Processing ID
12	8	ID	R	Version ID
13	15	NM	O	Sequence number
14	180	ST	O	Continuation pointer
15	2	ID	O	Accept acknowledgment type
16	2	ID	O	Application acknowledgment type
17	2	ID	O	Country code
18	6	ID	O	Character set
19	60	CE	O	Principal language of message

Practically, all information is entered by the user on primary setup of the application in the IHDMS. A representative result of the MSH message according to Table 3 is shown as below.

MSH||IHDMS|IHDMSMobilePhone|IHDMSMobilePhone|Lab|11101026914414||ORU^R10001075067129|P|2.4|1|||NE||ASCII|||

The next segment in the HL7 is the PID (Patient Identification) segment. This segment contains 30 different fields (see Table 4), such as patient ID number, sex, address, phone number, marital status, primary language, nationality, citizenship, and religion.

A representative result of the PID segment in a manner corresponding to Table 4 is shown as below.

PID||6804010117||Lee^Malrey||||||||||||0803738490||||||||||

The next segment after PID is the OBX (Observation) segment. The OBX segment is a part of multiple message types that transmit patient clinical information. 17 different fields of the OBX segment are detailed in Table 5.

A sample result of the OBX segment in a manner corresponding to Table 5 is shown as below.

OBX||ST|ECG Nanosensor||2.675^2.925^2.855^|ECG Nanosensor||||||||||

Each segment information of the MSH, PID, and OBX are mandatory to produce an accomplished HL7 message that can be delegated to the healthcare server. An individual HL7 message consists of MSH, PID, and OBX segments, which are produced by IHDMS. HL7 messages was transferred to the healthcare server afterward producing process.

Table 4: PID segment field information.

Seq.	Len.	Fmt	Opt.	Field name
0	3		R	Segment ID = "PID"
1	4	SI	O	Set ID—patient ID
2	12	CX	O	Patient ID (external ID)
3	16	CX	R	Patient ID (internal ID)
4	20	CX	O	Alternate patient ID—PID
5	48	PN	R	Patient name
6	48	PN	O	Mother's maiden name
7	14	TS	O	Date/time of birth
8	1	IS	O	Sex
9	48	PN	O	Patient alias
10	1	IS	O	Race
11	106	AD	O	Patient address
12	4	IS	O	County code
13	20	TN	O	Phone number—home
14	20	TN	O	Phone number—business
15	20	CE	O	Native language
16	1	IS	O	Marital status
17	3	IS	O	Religion
18	12	CX	R	Patient account number
19	11	ST	O	SSN number—patient
20	25	ST	O	Driver's license number—patient
21	9	CX	O	Mother's identifier
22	3	IS	O	Ethnic group
23	20	ST	O	Birth place
24	2	ID	O	Multiple birth indicator
25	2	NM	O	Birth order
26	4	IS	O	Citizenship
27	60	CE	O	Veterans military status
28	80	CE	O	Nationality
29	8	TS	O	Patient death date and time
30	1	ID	O	Patient death indicator

3.4. The IHDMS Android Application. Once the task assignment algorithm assigns the task to a particular nanosensor, then "Personal Server" is solely responsible for collecting data and events from the wireless body area network (WBAN). The "Personal Server" provides the user interface, controls the WBAN, fuses data and events, and creates unique session archive files. The software is implemented in Java for Android Mobile Operating System. It runs on Mobile Phone with Android OS. The Intelligent Healthcare Data Management System (IHDMS) Application was developed and implemented in this study. Android OS was used in our developed application. Figure 6 illustrated Start Page of the Android application for IHDMS.

A current user's health information can be seen by clicking button "View Information," if it is allowed already. At first time, it will be blank, because there is nobody, and any nanosensor was not allowed. User or caregiver permeates to the system through enter mandatory information by

Table 5: OBX segment field information.

Seq.	Len.	Fmt	Opt.	Field name
0	3		R	Segment ID = "OBX"
1	10	SI	O	Set ID—OBX
2	2	ID	R	Value type
3	20	CE	R	Observation identifier
4	20	ST	O	Observation Sub-ID
5	10	NM	R	Observation value
6	20	CE	R	Units
7	10	ST	O	References range
8	5	ID	O	Abnormal flags
9	5	NM	O	Probability
10	2	ID	O	Nature of abnormal test
11	1	ID	O	Result status
12	14	TS	O	Date of last normal values
13	20	ST	O	User defined access checks
14	14	TS	O	Date/time of the observation
15	60	CE	O	Producer's ID
16	80	CN	O	Responsible observer
17	60	CE	O	Observation method

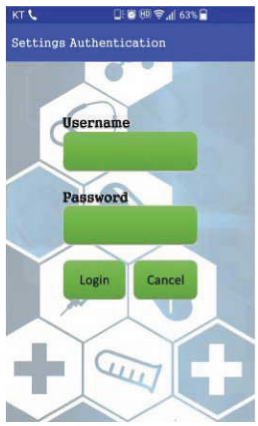


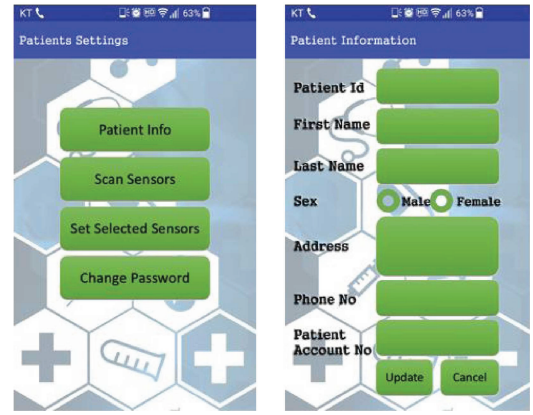
Figure 7: Settings Authentication page.

"Options" button. Figure 7 shows a Settings Authentication page by clicking "Options" button in Figure 8.

After updating the patient information, the current page Patient Information is redirected to the previous page of Patient Settings, as illustrated in Figure 9. Next step is to set the nanosensor information. A Page of the Set Selected Nanosensors in Android Application is illustrated in Figure 10. Figure 11 presents Nanosensor Settings Page and result of the ECG nanosensor.

Acquired nanosensor data is sent to micro gateway, and the micro gateway data is sent to mobile device. Mobile device application displays received information to the user. Figure 12 illustrated data receiving structure and result of the Blood pressure, oxygen saturation in the mobile device.

Figure 8: Start page of the Android application in Health Monitoring System.



(a) Patient Settings page (b) Patient Information page

Figure 9: Patient Settings and Patient Information pages.

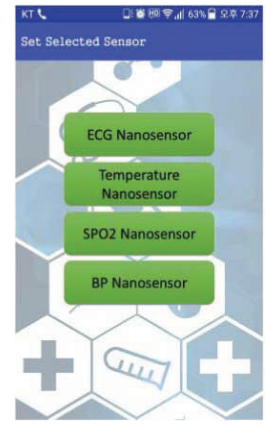
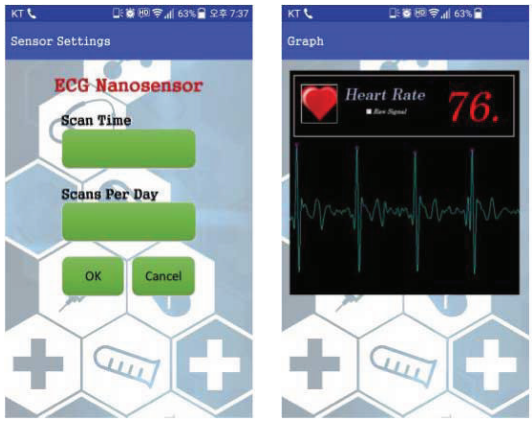


Figure 10: Pages of the set nanosensors and nanosensor settings.

4. Conclusions and Future Works

The proposed Intelligent Healthcare Data Management System dynamically assembles the nanosensor information of the sick person or patient and operates it on the smart phone.



(a) Nanosensor settings page of the EGG nanosensor

(b) Result page of the ECG nanosensor

Figure 11: Nanosensor settings and result page of the ECG nanosensor.

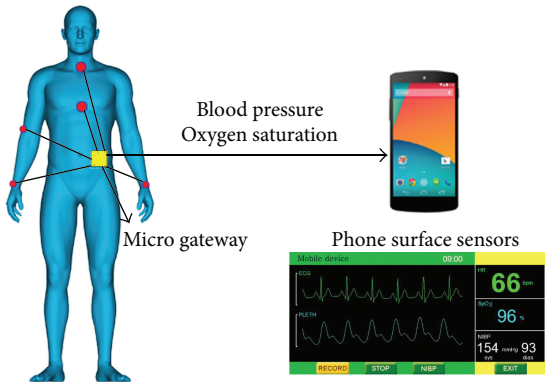


Figure 12: Data receiving structure and result of the Blood pressure, oxygen saturation in the mobile device.

It conducts the patient's health information in a superior prolific process.

This system also transforms raw data of the user from nanosensors to HL7 format and transfers the encrypted HL7 data to a remote healthcare server.

The main challenge faced by this system is in terms of the limited resources on the mobile environment. Similar to other applications operating in a mobile environment, the IHDMS could be deeply affected by context changes and scarcity of mobile platform's resource like network bandwidth, battery power, and computational power of handhelds.

This converted data dispatches to a server of its system. The battery lifetime of the facility is feasible to increase, the memory usage is less than 100 KB, and it operates all data by employing few and far between resources. Moreover, the proposed system decreases the waiting time in the transposing data, and secured channel was used for the server of the healthcare center in the running HL7 format data.

In the future, we will carry on this research in order to consummate the main point interfacing facilitator, to relate with other varying types of patient care devices that appear in medical industry. Also, the task assignment algorithm (TAA) can be amplified to utilize the motion nanosensor of the mobile device and apply this information in assigning task. We desire that the HL7 middleware library package will be maturing, to become embedded software that can be used in ubiquitous healthcare devices.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research is funded by research funds of Wonkwang University in 2016. This research was supported by Next-Generation Information Computing Development Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (2014M3C4A7030503). Also, this study was supported by the National Research Foundation of Korea (NRF) Grant funded by the Korea government (MSP) no. 2017R1A2B400667.

References

[1] A. Savić, D. Tošić, M. Marić, and J. Kratica, "Genetic algorithm approach for solving the task assignment problem," *Serdica Journal of Computing*, vol. 2, pp. 267–276, 2008.

[2] C.-H. Lee and K. G. Shin, "Optimal task assignment in homogeneous networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 8, no. 2, pp. 119–129, 1997.

[3] H. Mei, B. J. Van Beijnum, P. Pawar, I. Widya, and H. Hermens, "A*-based task assignment algorithm for context-aware mobile patient monitoring systems," in *2009 15th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications*, Beijing, China, 2009.

[4] B. Zhao, M. Wang, Z. Shao, J. Cao, K. C. C. Chan, and J. Su, "Topology aware task allocation and scheduling for real-time data fusion applications in networked embedded sensor systems," in *2008 14th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications*, pp. 293–302, Kaohsiung, Taiwan, 2008.

[5] P. Pawar, H. Mei, I. Widya, B.-J. van Beijnum, and A. van Halteren, "Context-aware task assignment in ubiquitous computing environment - a genetic algorithm based approach," in *2007 IEEE Congress on Evolutionary Computation*, pp. 2695–2702, Singapore, Singapore, 2007.

[6] R. B. Bachouch, A. Guinet, and S. Hajri-Gabouj, "An optimization model for task assignment in home health care," in *2010 IEEE Workshop on Health Care Management (WHCM)*, Venice, Italy, 2010.

[7] J. Hu and R. Marculescu, "Energy- and performance-aware mapping for regular NoC architectures," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 24, no. 4, pp. 551–562, 2005.

[8] W. Alsalih, S. Akl, and H. Hassanein, "Energy-aware task scheduling: towards enabling mobile computing over MANETs," in *19th IEEE International Parallel and Distributed Processing Symposium*, Denver, CO, USA, 2005.

[9] C. Otto, A. Milenkovic, C. Sanders, and E. Jovanov, "System architecture of a wireless body area sensor network for ubiquitous health monitoring," *Journal of Mobile Multimedia*, vol. 1, pp. 307–326, 2006.

[10] G. Paré, M. Jaana, and C. Sicotte, "Systematic review of home telemonitoring for chronic diseases: the evidence base," *Journal of the American Medical Informatics Association*, vol. 14, no. 3, pp. 269–277, 2007.

[11] J.-S. Kang, G. Jang, and M. Lee, "Stress status classification based on EEG signals," *The Journal of The Institute of Internet, Broadcasting and Communication*, vol. 16, no. 3, pp. 103–108, 2016.

[12] O. Sun-Jin, "Design of a smart application for remote diagnosis in ubiquitous computing environment," *The Journal of The Institute of Internet, Broadcasting and Communication*, vol. 16, no. 4, pp. 81–87, 2016.

[13] J. Kim, J. So, W. Choi, and K.-H. Kim, "Development of a mobile healthcare application based on tongue diagnosis," *The Journal of The Institute of Internet, Broadcasting and Communication*, vol. 16, no. 4, pp. 65–72, 2016.

[14] V. Kumar, B. Gupta, and S. K. Ramakuri, "Wireless body area networks towards empowering real-time healthcare monitoring: a survey," *International Journal of Sensor Networks*, vol. 22, no. 3, pp. 177–187, 2016.

[15] J. Hanen, Z. Kechaou, and M. B. Ayed, "An enhanced healthcare system in mobile cloud computing environment," *Vietnam Journal of Computer Science*, vol. 3, no. 4, pp. 267–277, 2016.

[16] N. Hamdi, R. Oweis, H. A. Zraiq, and D. A. Sammour, "An intelligent healthcare management system: a new approach in work-order prioritization for medical equipment maintenance requests," *Journal of Medical Systems*, vol. 36, no. 2, pp. 557–567, 2012.

[17] J. O'Donoghue and J. Herbert, "Data management within mHealth environments: patient sensors, mobile devices, and databases," *Journal of Data and Information Quality*, vol. 4, no. 1, article 5, pp. 1–20, 2012.

[18] M. Lee, T. M. Gatton, and K.-K. Lee, "A monitoring and advisory system for diabetes patient management using a rule-based method and KNN," *Sensors*, vol. 10, no. 4, pp. 3934–3953, 2010.

[19] D. Ketan, R. R. Bodhe, A. N. Sawant, and A. Kazi, "Proposed mobile based healthcare system for patient diagnosis using android OS," *International Journal of Computer Science and Mobile Computing*, vol. 3, no. 5, pp. 422–427, 2014.

[20] N. Lionel and J. Jong-Wook, "Addressing concurrency design for HealthCare Web Service Gateway in remote healthcare monitoring system," *International Journal of Advanced Smart Convergence*, vol. 5, no. 3, pp. 32–39, 2016.

[21] S. K. Sahoo, S. Parveen, and J. J. Panda, "The present and future of nanotechnology in human health care," *Nanomedicine: Nanotechnology, biology, and Medicine*, vol. 3, no. 1, pp. 20–31, 2007.

[22] E. Omanović-Mikličanin, M. Maksimović, and V. Vujović, "The future of healthcare: nanomedicine and internet of nano things," *Folia Medica Facultatis Medicinae Universitatis Saraeviensis*, vol. 50, no. 1, pp. 23–28, 2015.

[23] V. Raffa, O. Vittorio, C. Riggio, and A. Cuschieri, "*Progress in nanotechnology for healthcare*," *Minimally Invasive Therapy & Allied Technologies*, vol. 19, no. 3, pp. 127–135, 2010.

[24] N. A. Ali and M. Abu-Elkheir, "Internet of nano-things healthcare applications: requirements, opportunities, and challenges," in *2015 IEEE 11th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, Abu Dhabi, United Arab Emirates, 2015.

Hindawi
Journal of Sensors
Volume 2017, Article ID 1260734, 8 pages
https://doi.org/10.1155/2017/1260734

Research Article

Detection of Freezing of Gait Using Template-Matching-Based Approaches

Cheng Xu,[1,2] Jie He,[1,2] Xiaotong Zhang,[1,2] Cunda Wang,[1,2] and Shihong Duan[1,2]

[1]*School of Computer and Communication Engineering, University of Science and Technology, Beijing, China*
[2]*Beijing Key Laboratory of Knowledge Engineering for Materials Science, Beijing, China*

Correspondence should be addressed to Jie He; hejie@ustb.edu.cn

Received 21 July 2017; Accepted 2 October 2017; Published 6 November 2017

Academic Editor: Mucheol Kim

Every year, injuries associated with fall incidences cause lots of human suffering and assets loss for Parkinson's disease (PD) patients. Thereinto, freezing of gait (FOG), which is one of the most common symptoms of PD, is quite responsible for most incidents. Although lots of researches have been done on characterized analysis and detection methods of FOG, large room for improvement still exists in the high accuracy and high efficiency examination of FOG. In view of the above requirements, this paper presents a template-matching-based improved subsequence Dynamic Time Warping (IsDTW) method, and experimental tests were carried out on typical open source datasets. Results show that, compared with traditional template-matching and statistical learning methods, proposed IsDTW not only embodies higher experimental accuracy (92%) but also has a significant runtime efficiency. By contrast, IsDTW is far more available in real-time practice applications.

1. Introduction

Parkinson's disease is a kind of common neurological disorder caused by dopamine and gradually loss of function of other subcortical neurons. PD usually causes the patients' movement function disorder, starting from tremors of one side body or activity clumsy, and further involves the contralateral limb [1, 2]. Clinical manifestations of Parkinson's disease are mainly for static tremor, bradykinesia, myotonia, and freezing of gait. Among them, FOG is a kind of typical symptom. The patient is not easily maintaining the balance of the body and is likely to fall on the road surface with even a bit uneven. Its typical symptoms are loss of ability to walk in a sudden feet stuckness on the ground and disability to move in a few minutes or being no longer able to move again. FOG seems to be common in the start period of walking, turning, and moving close to the target or when one is worried whether he is able to get through the known obstacles, for example, getting through the revolving door. Every year, fall incidence rates range from 50% to 70%. It is one of the main reasons for disability to PD patients [1].

Freezing of gait (FOG) is one of the cardinal symptoms of the PD which is defined as an inability of a person to move one's feet in spite of the fact that he/she intends to move [3]. Existing methods for prevention and cure of FOG mainly rely on drugs, the most widely used of which is levodopa (LD) [4]. However, drug's effect duration is generally 2–6 h, and different patients' drug resistance is various to different kinds of drugs. These specificities may lead to large fluctuations, showing up as the patient suddenly cannot move or can suddenly move freely. These two appearances alternate in a few minutes, namely, "On-Off" phenomenon. Once "On-Off" appears, it is hardly to be cured [5].

In addition, some nondrug therapy methods can also be used in the prevention and treatment of FOG. PD patients complete the corresponding action according to the instructions (like music rhythm, visual cues, etc.) [6–8]. These coordination practices are proved to be effective in keeping FOG from getting worse. Plotnik et al. [6] suggested that external tempo clues were of great help to increase walking speed significantly for PD patients. Compared with treatment methods relying on drugs, with the instruction of external information, PD patients could be prepared in advance and can respond to the possibility of FOG, as well as the happening of "On-Off" phenomenon.

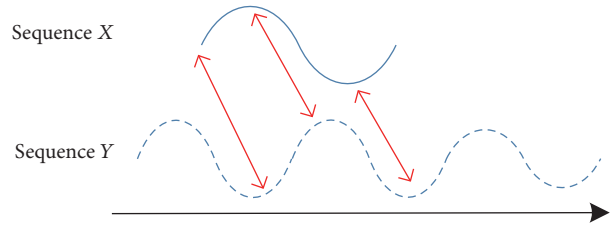


Figure 1: A typical pair of sequences for template matching.

For both drug and nondrug therapy methods, the detection and forewarning of FOG are significantly important. Wearable sensors are widely used to realize the real-time detection and alarming of FOG for PD patients. Previous studies usually focus on the fusion of sensors like baroreceptor, IMU, and so on. They capture the sensor signals when PD patients in activity do deep analysis of the wavelets and give out alarms before FOG occurs. The research topics about forewarning of FOG based on wearable sensors mainly concentrate on the selection of sensors [5], locations [9], and high effective algorithms [10].

Different from the traditional statistical methods [5–10], template matching is a high effective recognition method with both high recognition accuracy and efficiency, which has been applied to physical activity [11–13]. Muscillo et al. [11] proposed user-dependent templates to target recognition of arm-specific tasks. Likewise, Chen and Shen [12] focused on recognizing activities performed with the right upper limb using a classification framework based on template matching. Stiefmeier et al. [13] proposed an innovative approach consisting of encoding motion data into sequence of finite symbols and performing activity recognition by using string-matching algorithms. However, to the best of our known, no effort has been paid in the real-time detection of FOG.

To summarize, traditional statistical methods for FOG detection have low accuracy and efficiency, and it can hardly meet the requirements for practice real-time applications. Template-matching method is of high performance advantages; however, it is seldom used in FOG detection. The purpose of this paper was to investigate the use of template matching for the detection of FOG.

The rest of this paper is organized as follows: typical template-matching methods are introduced in Section 2. In Section 3, the framework of our system is presented and an improved sDTW algorithm (IsDTW) is illustrated. The proposed algorithm is verified on an open dataset and results are analyzed in Section 4. Proposed IsDTW algorithm is compared with both traditional template methods and statistical methods in recognition accuracy and real-time performance. The conclusions of our work are given in Section 5.

2. Template-Matching Methods

Template-matching algorithm is an approach for comparing two-time sequences in terms of both their state and dynamics. Time series can be used to classify primitive physical activities from data provided by wearable sensors, such as accelerometers. Template matching is a high-level machine learning technique that identifies the parts on one sample that match a predefined template.

2.1. Problem Formulation. A typical pair of sequences for template matching is shown in Figure 1. Define sequence X as template signal and Y as recorded signal, whose lengths are, respectively, n and m; namely,

$$\begin{aligned} X &= (t_1, x_1), (t_2, x_2), \ldots, (t_i, x_i), \ldots, (t_n, x_n), \\ Y &= (t_1, y_1), (t_2, y_2), \ldots, (t_j, y_j), \ldots, (t_m, y_m). \end{aligned} \tag{1}$$

Each subsequence of the time series will be compared with the template X. Distance based similarity between the template and the time series segment (denoted as $\mathrm{dist}(X, Y_{ij})$) is computed. Finally, we will identify the best similarity measures by the distance as the probable template occurrences. The smaller the distance is, the more similar they are. We are intending to find the optimal subsequence in Y whose $\mathrm{dist}(X, Y_{ij})$ value is smaller than given threshold.

Distance based similarity measurement method could be various, and we will introduce some of the most widely used methods.

2.2. Typical Template-Matching Methods

2.2.1. Euclidean Distance. Euclidean metric, namely, Euclidean distance, is a common adopted definition of distance. It refers to the actual distance between two points in multidimensional space or the natural length of vector (namely, the distance between this point and the origin point). Respectively, denote $X = [x_1, x_2, \ldots, x_i, \ldots, x_m]$ and $Y = [y_1, y_2, \ldots, y_i, \ldots, y_n]$ as two temporal sequences. Thus, the distance d_i ($i = 0, \ldots, n - m - 1$, $n < m$) could be calculated from vectors X and Y. For the ith sample, the regularized Euclidean distance could be represented as

$$d_i = \sqrt{\sum_{k=1}^{m} (Y(i+k) - X(k))^2}. \tag{2}$$

2.2.2. Dynamic Time Wrapping (DTW). DTW could be used to measure the similarity or distance of these two sequences. The core of DTW is based on the idea of dynamic programming (DP), automatically searching for the optimal

path with local optimization method. Taking the minimum accumulation of distortion between two vectors as the objective could avoid errors caused by different time length.

In order to align the sequences X and Y, a $n \times m$ matrix is needed, while element (i, j) represents the DTW distance $\|d(x_i, y_j)\|$ (generally 1st normal form $\|d(x_i, y_j)\|_1$) between x_i and y_j. Namely, each element in the matrix stands for the similarity between two points in X and Y, and the smaller the distance is, the greater the similarity is. DP is applied to find optimal path crossing a number of grids in the matrix and calculation is conducted among the points crossed by the path.

The paths that satisfy all above conditions could be as many as exponential, but the minimum cost path is the one we are interested in. Thus, the following equation could be achieved:

$$\text{Dist}(X, Y) = d\left(x_i, y_j\right) + \min \begin{cases} \text{Dist}\left(i-1, j-1\right) \\ \text{Dist}\left(i-1, j\right) \\ \text{Dist}\left(i, j-1\right). \end{cases} \tag{3}$$

2.2.3. Subsequence Dynamic Time Wrapping (sDTW). sDTW is designed for searching repeated "child segments" from a long sequence. The core idea of sDTW is dividing distance matrix D into subbands and using traditional DTW to search the optimal path in subbands. Firstly, divide D into several inclined strip-like regions with the same width. Among the overlap, $s1$ and $s2$ are, respectively, the starting point of the two subbands. Assume that the displacement from $s1$ to $s2$ is R and the width of inclined region is $2R + 1$. Thus, for a $m \times n$ matrix, the number of its contained regions is $\lfloor (n-1)/R + (m-1)/R \rfloor$.

Afterwards, find the optimal path in each strip-like region using DTW. In each optimal path, only a small segment is corresponding to the similar parts of these two consecutive sequences. So we need to cut out the specific subpaths, which should meet these requirements: (1) the points contained in subpath; namely, the length of subpath is smaller than L; (2) the average of all points in the subpath; namely, the average of subpath is smaller than θ. Given a subpath with N points, whose length is L and its average is θ, working out the LCMA (length-constrained minimum average) is as follows:

$$f = \min_{1 \le s \le t \le N} \frac{1}{t-s+1} \sum_{k=s}^{t} \text{Dist}\left(i_k, j_k\right), \quad t - s + 1 \ge L. \tag{4}$$

2.2.4. Cross Correlation. In signal processing, it is often to study the similarity of two signals, in order to implement signal detection, recognition, and extraction. The method that could be used to analyze the similarity of signals is called cross correlation. Given two temporal sequences X and Y, whose length is, respectively, n and m, its cross correlation function is defined as follows:

$$C_{YX}(\tau) = \frac{1}{n-1} \sum_{i=0}^{n-1} [Y(i+\tau)][X(i)]. \tag{5}$$

Generally speaking, cross correlation index could be used to normalize the standard deviation of two signals, and the cross correlation coefficient is defined as follows:

$$\gamma_{YX}(\tau) = \frac{C_{YX}(\tau)}{\sigma_{YY}\sigma_{XX}}, \tag{6}$$

where σ_{XX} and σ_{YY} are, respectively, the standard deviation of X and Y and the value of $\gamma_{YX}(\tau)$ is between -1 and $+1$. If $\gamma_{YX}(\tau) = -1$, it illustrates that X and Y have the same shape but opposite phase. If $\gamma_{YX}(\tau) = 0$, it illustrates that X and Y have no similarities. If $\gamma_{YX}(\tau) = -1$, it illustrates that X and Y are totally the same. When the signal is compared with itself, it is called self-correlation, defined as follows:

$$\widehat{R}_{YY}(\tau) = \frac{1}{m-1} \sum_{i=0}^{m-\tau-1} [Y(i+\tau)][Y(i)]. \tag{7}$$

This function is often used to identify periodic signals from white noise in order to recognize signal cycle and repetitive patterns.

3. System Overview and Algorithm Design

Former researchers have studied the detection method for FOG with time and frequency domain features extracted by FFT [9]. These studies have already achieved a good accuracy but poor real-time characteristic. In this chapter, based on the point of template matching, we proposed an improved subsequence Dynamic Time Wrapping (IsDTW) method, to realize the real-time and high precision FOG detection and alarm. IsDTW gives out a good real-time performance as well as high accuracy.

The whole process could be divided into two stages:

(1) Data preprocessing: the main work is template generating and threshold confirmation, namely, model training stage.

(2) Subsequence searching: estimate the similarity using proposed algorithm, and detect FOG.

Notations. The original signal is denoted as $X = \{(t_1, x_1), (t_2, x_2), \ldots, (t_i, x_i), \ldots, (t_n, x_n)\}$, and the query subsequence is denoted as $Y = \{(t_1, y_1), (t_2, y_2), \ldots, (t_j, y_j), \ldots, (t_m, y_m)\}$. $X_{i,j}$ stands for the subsequence from time i to time j in sequence X. The framework of the whole system is presented in Figure 2.

3.1. Template Generation. The template sequence is $Y = \{(t_1, y_1), (t_2, y_2), \ldots, (t_j, y_j), \ldots, (t_m, y_m)\}$, and its length m is the predefined window length. The sequence contains several sensor data subsequences, such as accelerometer data A_x, A_y, A_z. The sample data contains lots of labeled FOG and non-FOG data. Divide these data into subsequences with length of m, and average them with method of "interp1" integrated in Matlab. Apply this operation to each axis of the subsequences, and finally the template of FOG is achieved, as shown in Figure 3.

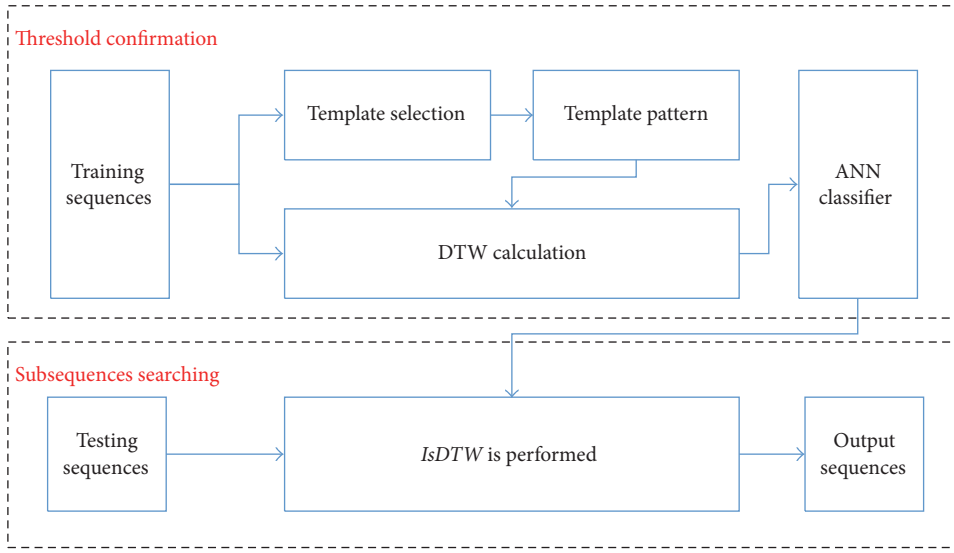


Figure 2: System framework diagram.

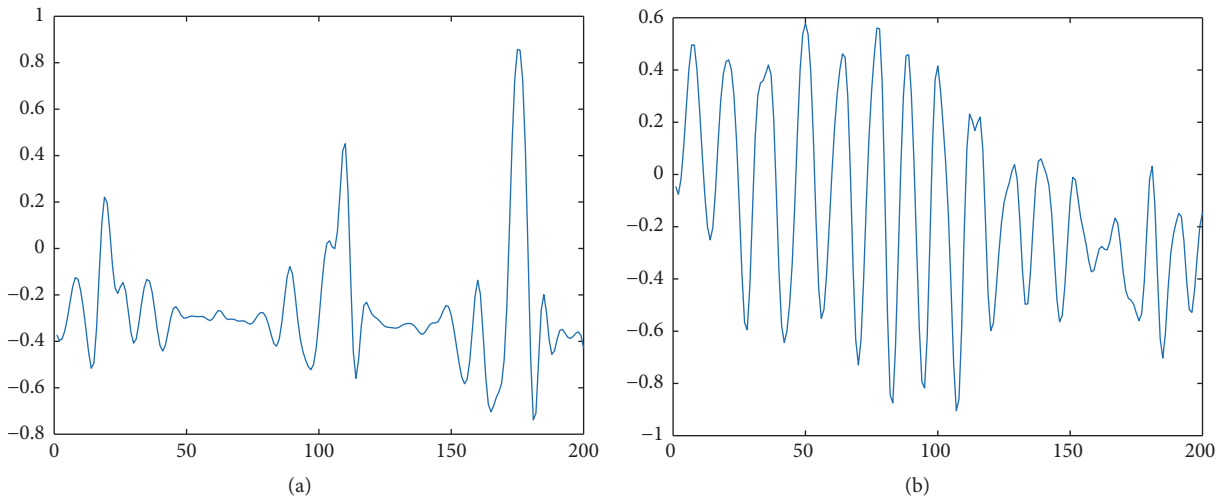


Figure 3: Gait template: (a) regular gaits; (b) FOG template.

3.2. Threshold Confirmation. The most urgent issue to be solved in proposed algorithm is how to determine the threshold value of loss of function. In common sense, due to the differences of various subjects and noise, the expected threshold value could vary from person to person. For the above reasons, in this paper, we proposed a dynamic threshold estimation method based on former statistical model. In data preprocessing stage, we construct an Artificial Neural Network (ANN) classifier to obtain the optimal threshold value ε, which is suitable for each subject, using labeled training data. This method depends on machine learning technique, running on a large amount of actual data, which makes it have higher running speed and credibility.

3.3. Similarity Computation: Improved Subsequence DTW: IsDTW. Similar to DTW, we proposed this improved sDTW method to compute the similarity of two sequences by updating distance matrix. In each loop of the algorithm processing, two variables are stored, $D(t,k)$ and $X(t,k)$. $D(t,k)$ denotes the minimum DTW value of sequence Y and subsequence $S_{i,t}$. $X(t,k)$ denotes the start time of sequence $X_{i,t}$, namely, $i = X(t,k)$. $D(t,k)$ could be obtained by following methods:

$$D(t,k) = \|x_t - y_k\| + D_{\text{best}},$$

$$D_{\text{best}} = \min \begin{cases} \text{Dist}(t,k-1) \\ \text{Dist}(t-1,k) \\ \text{Dist}(t-1,k-1), \end{cases} \tag{8}$$

Input: x_t at time t
Output: similar subsequences: $S_{i,t}$
For $k = 1$ to m **do**
  Calculate $D(t,k)$ and $X(t,k)$
End
IF $D_{\min} \le \varepsilon$ **then**
  IF $\forall D(t,k) > D_{\min} \vee X(t,k) > t_e$ **then**
    Return $D_{\min}, t_s, t_e$
    $D_{\min} = +\infty$
    For $k = 1$ to m **do**
      IF $X(t,k) \le t_e$ **then**
        $D(t,k) = +\infty$
      End
    End
  End
End
If $D(t,m) \le \varepsilon \wedge D(t,m) \le D_{\min}$ **then**
  $D_{\min} = D(t,m)$
  $t_s = X(t,m)$
  $t_e = t$
End
For $k = 0$ to m **do**
  $D(t-1,k) = D(t,k)$
  $X(t-1,k) = D(t,k)$
End

Algorithm 1

where $D(t,0) = 0$ and $D(0,0) = D(0,k) = +\infty (t = 1,2,\ldots,n$ and $k = 1,2,\ldots,m)$. Similarly, $X(t,k)$ could be obtained by following methods:

$$X(t,k) = \begin{cases} X(t,k-1), & \text{If } D_{\text{best}} = D(t,k-1) \\ X(t-1,k), & \text{If } D_{\text{best}} = D(t-1,k) \\ X(t-1,k-1), & \text{If } D_{\text{best}} = D(t-1,k-1), \end{cases} \quad (9)$$

where $X(t,0) = t$.

IsDTW is targeted for searching all possible subsequences $X_{i,j}$ that the similarity satisfies the given threshold ε between sequence X and template Y, namely $\text{Dist}(X_{i,j}, Y) \le \varepsilon$, where $j = i + m - 1$ and $i = 1,2,\ldots,n-m+1$. The whole algorithm could be described as shown in Algorithm 1.

4. Results and Analysis

This section may be divided by subheadings. It should provide a concise and precise description of the experimental results and their interpretation as well as the experimental conclusions that can be drawn.

4.1. Datasets. The method proposed in this paper is verified on open source dataset [14]. Bachlin et al. recruited 10 PD patients as experimental subject, among which there are 7 male and 3 female, whose ages are 66.5 ± 4.8 years and disease duration is 13.7 ± 9.67 years. All subjects are with the history of FOG and could walk freely in the condition of "off-medicine" without external assistance. All data is collected and analyzed in the condition of "off-medicine."

Subjects are required to complete the data collecting under the following experimental scenarios: walking forward and backward in a straight line, randomly walking and stopping, rotating 360 degrees, and daily life activities. The entire experimental processes are recorded by video camera. Two medical personnel diagnose the two possible conditions, FOG and non-FOG, from the real-time video information. Every time FOG occurs, record the starting and ending time. Each subject is mounted with three triaxial accelerometers, respectively, placed on shank, thigh, and lower back. The sampling rate is set as 64 Hz. Eventually, sampled from 10 PD subjects, 8 hours and 20 minutes data are collected, which contains 237 events of FOG.

Figure 4 shows one patient's actual recorded data during the experimental process. The subject ran into FOG condition exactly at the very beginning of the whole process and then continued walking normally after a short break; then, sensors detect trembling caused by FOG, and the legs got stuck; after a long time of FOG phenomenon, the subject came into a very fast march; finally, the subject rested down after a short FOG. The experimental results are matched with those given by the two medical personnel, which are marked in the figure with color fonts. Moreover, as shown in the figure, in the first and third FOG segments, short FOG stuckness is recognized rightly other than mistaken for rest; in the second FOG segment, the leg trembling caused by FOG is also recognized. These two points further verify the effectiveness of the proposed algorithm.

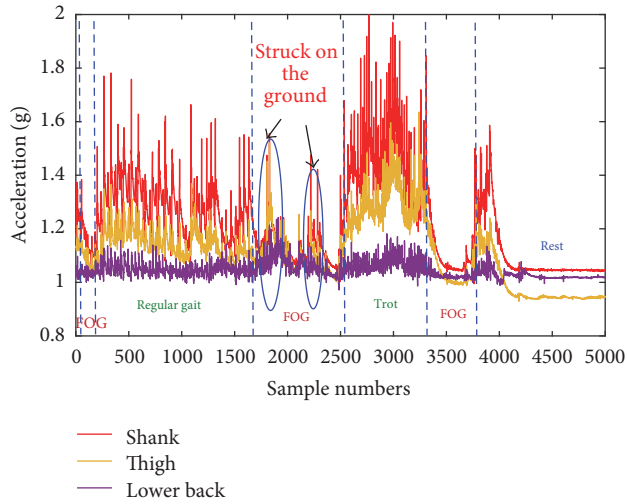


Figure 4: A typical pair of sequences for DTW comparison.

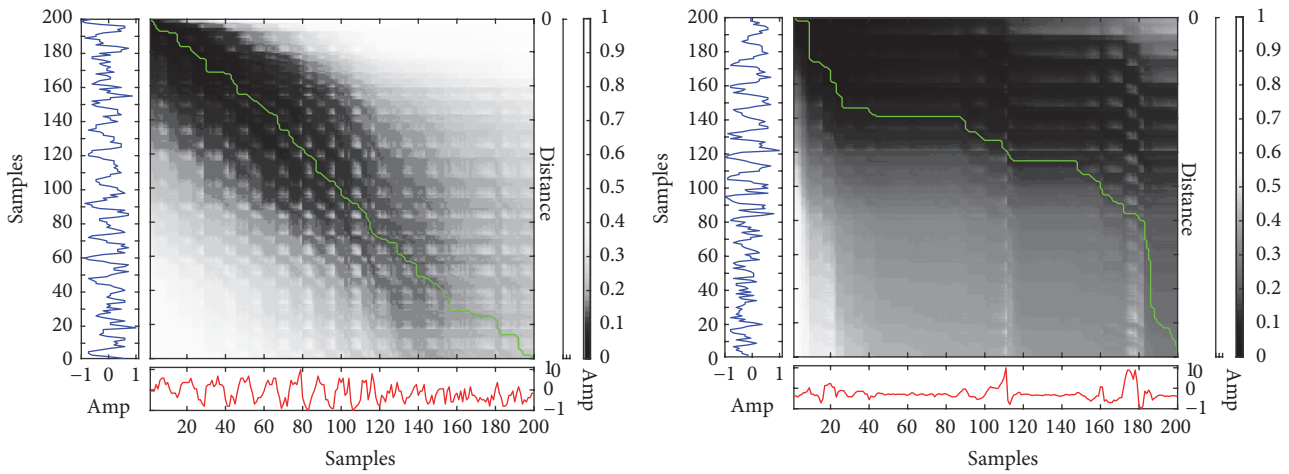


Figure 5: IsDTW matching diagram: (a) with FOG; (b) with regular gaits.

4.2. Results Discussion. Figure 5 shows the template-matching results with using IsDTW to detect FOG and regular gaits. It could be clearly seen that when FOG template compares with pathological data, shown in Figure 5(a), the DTW path tends to be more straight and the cumulative distance (namely, D) is smaller. It indicates that the compared two subsequences have higher similarity, and in other words, the detection result is FOG. When FOG template compares with disease-free data, shown in Figure 5(b), the DTW path is more winding and the cumulative distance is bigger, and the detection result is non-FOG.

The effect of FOG detection and sensor locations has been analyzed separately by ROC curve as follows. IsDTW is applied to the dataset and ROC curves of each subset are drawn in Figure 6, namely, lower back mounted sensor (blue curve), thigh mounted sensor (brown curve), shank mounted sensor (red curve), and all three sensors (green curve). Since the detection method has a better performance when the ROC curve is closer to upper left corner, we conclude from visual inspection that it is with better performance when all sensor data is used. When only shank data is adopted, the result shows less better performance. This may be because the shank part can more reflect the characteristics of FOG.

4.3. Compared with Template-Matching Methods. In this paper, several template-matching methods are selected to do FOG detection on open source dataset [14]. Various experimental results are achieved, as shown in Figure 7. The overall dataset contains sensor data collected from three accelerometers, shank, thigh, and lower back. Experiments are conducted with the whole dataset and each part of it, respectively. In different scenarios, our proposed IsDTW algorithm all outperformed the others. Following conclusions could be achieved:

(1) Each algorithm performs differently on various datasets. The best result was obtained when all data

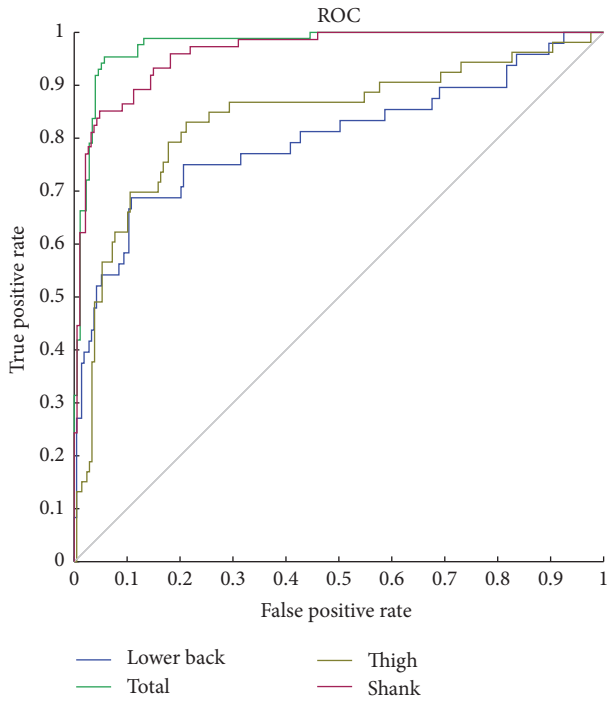


Figure 6: ROC diagram.

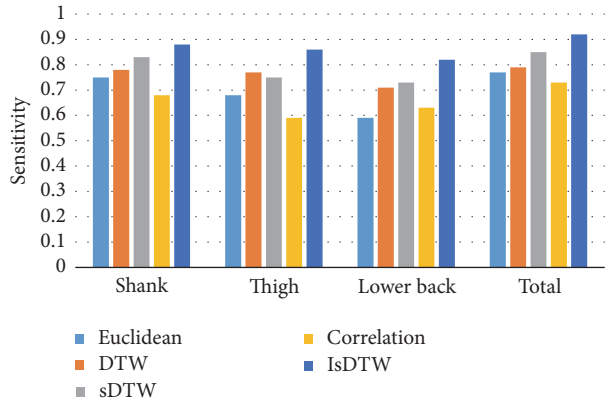


Figure 7: Results comparison of IsDTW on different datasets.

is utilized, while the second best is with only shank data. Thus, it can be seen that shank responses to the FOG phenomenon for maximum efficiency, and the other parts play a supporting roles.

(2) On different datasets, the proposed IsDTW algorithm all achieved the best result, which can to some extent reflects the stability of the algorithm. Meanwhile, the best result could be as high as 92%. Compared with the others, IsDTW keeps updating the minimum distance path and maintain a high efficiency as well as high accuracy. Besides, using ANN to update the dynamic threshold, it improves the migration ability of the algorithm.

Table 1: The list of selected features.

Features	Domain
Mean	Time
Variance	Time
Root mean square	Time
Range (maximum–minimum)	Time
Total energy	Time
Skewness	Time
Main frequency	Frequency
Entropy	Frequency
Quartile	Frequency

(3) Generally, DTW related algorithms work better than Euclidean and cross correlation, which indicates that DTW has more advantages for the identification of sequences with variable length.

4.4. Compared with Statistical Methods. For further verification of proposed IsDTW, comparison experiments are carried out between template-matching FOG detection methods and traditional statistical classification methods, including Decision Tree (DT), Naïve Bayesian Network (NBN), and Artificial Neural Network (ANN). According to [15], 13 common used time-domain and frequency domain features are selected, and 9 of them are chosen to be applied to classification with Relief method [16]. The chosen features are listed in Table 1.

Taking advantage of the open source tool Weka [17], comparison experiments are all performed and compared with proposed template-matching method. Results are shown in Table 2. It can be seen that IsDTW has significant advantages in classification accuracy, namely, IsDTW > ANN > NBN > DT. ANN stands out among statistical methods with an accuracy 0.88 but is still lower than that of IsDTW. In the meantime, its runtime efficiency performs lower than that of IsDTW which can hardly meet the requirements of real-time in practice applications. As DT is a kind of lightweight algorithm with simple principle, it has high efficiency but its accuracy may not meet the real-time needs.

Therefore, the IsDTW method proposed in this paper has the advantages of both high precision and real-time capability, and it may meet the demands of practical application.

5. Conclusions

In this paper, we discuss the detection of FOG with utilizing of template-matching methods. Contrast experiments are carried out on open source dataset OPPORTUNITY. Template-matching methods are compared with Euclidean, DTW, sDTW, and cross correlation. Experimental results show that template-matching methods have certain advantages, and our proposed IsDTW apparently has higher accuracy. For comparison, IsDTW is compared with nontemplate methods (statistical methods), and the results show that our algorithm has not only higher experimental accuracy but to a

Table 2: Performance comparison of IsDTW with statistical machine learning methods.

Methods	DT	NBN	ANN	IsDTW
Accuracy	0.77	0.81	0.88	**0.92**
Runtime (s)	0.55	1.26	3.52	0.64

certain extent is better than traditional methods on runtime efficiency, making it more applicable in practice applications.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Cheng Xu and Jie He are co-first authors and contributed equally to this work.

Acknowledgments

This work is supported by National Natural Science Foundation of China (NSFC), Projects no. 61671056, no. 61302065, no. 61304257, and no. 61402033, the National Key R&D Program of China, no. 2016YFC0901303, Beijing Natural Science Foundation, Project no. 4152036, and Tianjin Special Program for Science and Technology, no. 16ZXCXSF00150.

References

[1] N. E. Allen, A. K. Schwarzel, and C. G. Canning, "Recurrent falls in Parkinson's disease: a systematic review," *Parkinson's Disease*, vol. 2013, Article ID 906274, 16 pages, 2013.

[2] J. A. Obeso, C. W. Olanow, and J. G. Nutt, "Levodopa motor complications in Parkinson's disease," *Trends in Neurosciences*, vol. 23, no. 10, pp. S2–S7, 2000.

[3] J. G. Nutt, B. R. Bloem, N. Giladi, M. Hallett, F. B. Horak, and A. Nieuwboer, "Freezing of gait: moving forward on a mysterious clinical phenomenon," *The Lancet Neurology*, vol. 10, no. 8, pp. 734–744, 2011.

[4] F. C. F. Chang, D. S. Tsui, N. Mahant et al., "24h Levodopa-carbidopa intestinal gel may reduce falls and "unresponsive" freezing of gait in Parkinson's disease," *Parkinsonism & Related Disorders*, vol. 21, no. 3, pp. 317–320, 2015.

[5] S. T. Moore, D. A. Yungher, T. R. Morris et al., "Autonomous identification of freezing of gait in Parkinson's disease from lower-body segmental accelerometry," *Journal of NeuroEngineering and Rehabilitation*, vol. 10, no. 1, article 19, 2013.

[6] M. Plotnik, S. Shema, M. Dorfman et al., "A motor learning-based intervention to ameliorate freezing of gait in subjects with Parkinson's disease," *Journal of Neurology*, vol. 261, no. 7, pp. 1329–1339, 2014.

[7] P. Arias and J. Cudeiro, "Effect of rhythmic auditory stimulation on gait in parkinsonian patients with and without freezing of Gait," *PLoS ONE*, vol. 5, no. 3, Article ID e9675, 2010.

[8] P. Arias and J. Cudeiro, "Effects of rhythmic sensory stimulation (auditory, visual) on gait in Parkinson's disease patients," *Experimental Brain Research*, vol. 186, no. 4, pp. 589–601, 2008.

[9] F. B. Horak and M. Mancini, "Objective biomarkers of balance and gait for Parkinson's disease using body-worn sensors," *Movement Disorders*, vol. 28, no. 11, pp. 1544–1551, 2013.

[10] A. Muro-de-la-Herran, B. García-Zapirain, and A. Méndez-Zorrilla, "Gait analysis methods: an overview of wearable and non-wearable systems, highlighting clinical applications," *Sensors*, vol. 14, no. 2, pp. 3362–3394, 2014.

[11] R. Muscillo, M. Schmid, S. Conforto, and T. DprimeAlessio, "Early recognition of upper limb motor tasks through accelerometers: Real-time implementation of a DTW-based algorithm," *Computers in Biology and Medicine*, vol. 41, no. 3, pp. 164–172, 2011.

[12] C. Chen and H. Shen, "A feature evaluation method for template matching in daily activity recognition," in *Proceedings of the IEEE International Conference on Signal Processing, Communications and Computing, (ICSPCC '13)*, IEEE, China, August 2013.

[13] T. Stiefmeier, D. Roggen, and G. Tröster, "Fusion of string-matched templates for continuous activity recognition," in *Proceedings of the 11th IEEE International Symposium on Wearable Computers, ISWC 2007*, pp. 41–44, usa, October 2007.

[14] M. Chlin, M. Plotnik, D. Roggen et al., "Wearable assistant for Parkinson's disease patients with the freezing of gait symptom," *IEEE Transactions on Information Technology in Biomedicine A Publication of the IEEE Engineering in Medicine Biology Society*, vol. 14, no. 2, pp. 436–446, 2010.

[15] J. Margarito, R. Helaoui, A. M. Bianchi, F. Sartor, and A. G. Bonomi, "User-independent recognition of sports activities from a single wrist-worn accelerometer: a template matching based approach," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 4, pp. 788–796, 2016.

[16] Y. Sun, "Iterative relief for feature weighting: algorithms, theories, and applications," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 29, no. 6, pp. 1035–1051, 2007.

[17] I. H. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques With Java Implementations*, Morgan Kaufmann, Burlington, Mass, USA, 2002.