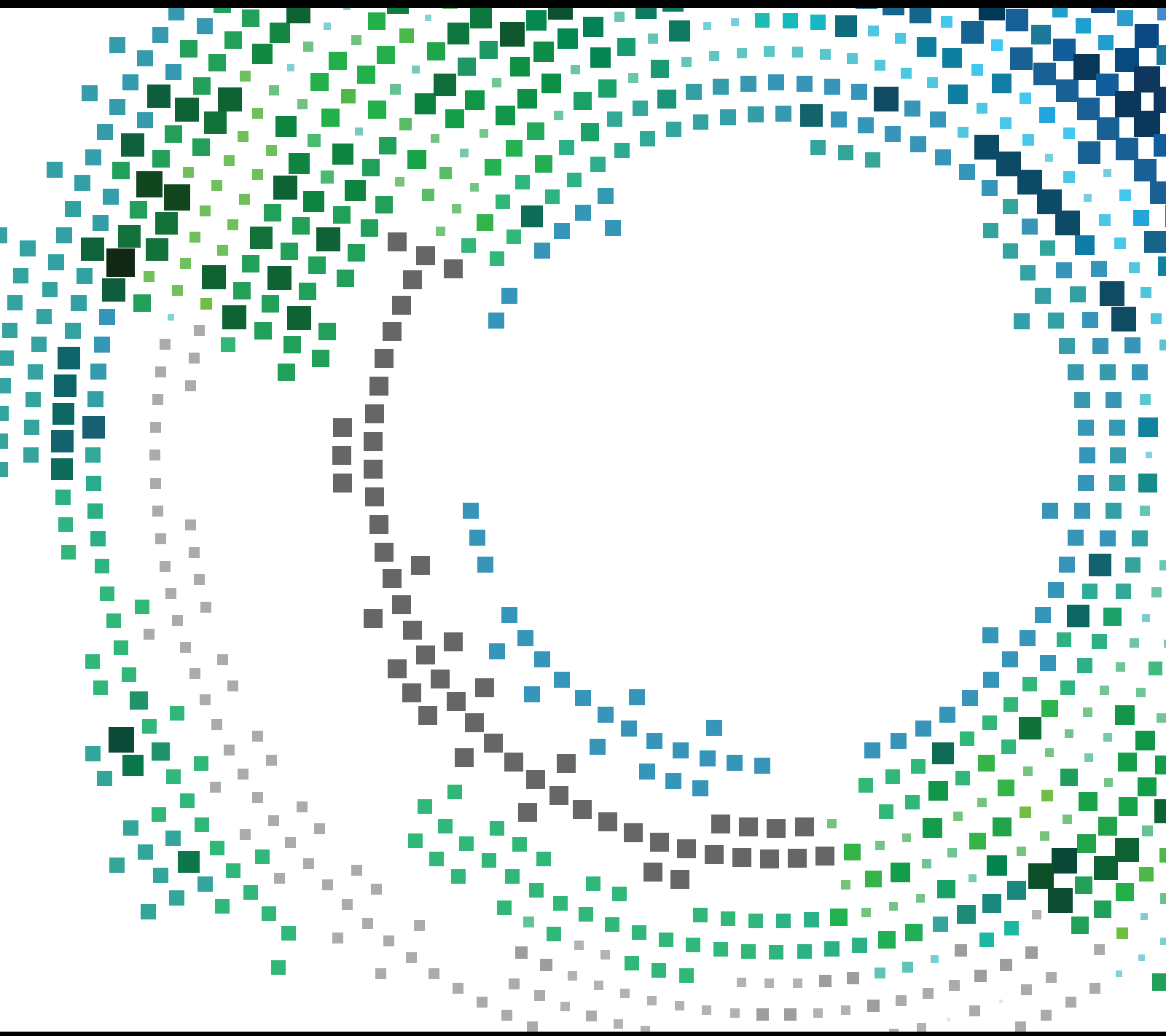
Mobile Information Systems

Fundamental Issues in Mobile Healthcare Information Systems

Guest Editors: Mehmat Orgun, Basit Shahzad, and Christoph Thuemmler



Fundamental Issues in Mobile Healthcare Information Systems

Mobile Information Systems

Fundamental Issues in Mobile Healthcare Information Systems

Guest Editors: Mehmat Orgun, Basit Shahzad,
and Christoph Thuemmler

Contents

Hindawi Publishing Corporation
Mobile Information Systems
Volume 2016, Article ID 6504641, 2 pages
http://dx.doi.org/10.1155/2016/6504641

Editorial

Fundamental Issues in Mobile Healthcare Information Systems

Basit Shahzad,[1] Mehmet A. Orgun,[2,3] and Christoph Thuemmler[4]

[1]*College of Computer & Information Science, King Saud University, Riyadh 12372, Saudi Arabia*
[2]*Intelligent Systems Group (ISG), Department of Computing, Macquarie University, Sydney, NSW 2109, Australia*
[3]*Faculty of Information Technology, Macau University of Science and Technology, Avenida Wai Long, Taipa 999078, Macau*
[4]*Edinburgh Napier University, Edinburgh EH10 5DT, UK*

Correspondence should be addressed to Mehmet A. Orgun; mehmet.orgun@mq.edu.au

Received 25 July 2016; Accepted 25 July 2016

1. Introduction

Consistent growth in the mobile technology has led to many significant features that have improved the quality of services provided to people from all walks of life. While the need for a broader and more diverse range of available services is growing, the multifaceted and deep rooted challenges of the mobile technology such as security, privacy, efficiency, coherence, and resilience have to evolve. While most of the online businesses, social networking, financial transactions, personal record managements, and so forth are increasingly being done from mobile phones, it is important that an appropriate infrastructure capable of handling this data is in place to keep, handle, update, secure, and make it available when needed in accordance with national and international rules and regulations. The quality of life can be improved by automating a number of tasks that have a direct impact on day-to-day living. One prominent area of growing interest in this regard is the support of healthcare provision anywhere, anyhow, and at any time. Suitable information systems and the relevant network infrastructures are moving closer and closer together. There is the need to have security, privacy, efficiency, robustness, consistency, and availability at all times. In order to incorporate the said objectives, it is justified to advocate an information system that can operate on mobile devices to provide healthcare services, whereby the service layer and the data transportation layer are moving increasingly closer to each other.

Healthcare systems are at the cusp of being revolutionized by advancements in technology, which when appropriately integrated into existing best practices can enable faster and safer cure, improved doctor-patient relationships, personalized treatments, and lower costs. Typically this can be measured and controlled by monitoring the Quality of Experience (QoE).

With rapid advances in computing and associated technologies, we are also seeing steady and seamless integration of communication, networking, hardware miniaturization, sensing, cryptography, and a range of algorithmic advances for smarter and increasingly personalized healthcare. At the forefront of challenges in emergent smarter healthcare systems the issues are how to manage massively growing amount of healthcare information and smart devices over a variety of technologies across different domains maintaining a guaranteed quality of service (QoS) at any time.

In the realm of healthcare, while there are clear opportunities to leverage information management emanating from today's computing technologies, additional challenges include providing information reliability, security, patient's privacy, real-time criticality, information fusion, system sustainability, and social interaction, among others. Although research in the domain of information management analytics for smarter healthcare is attracting attention across disciplines, critical applications, new opportunities, challenges, models, and technologies are yet to be explored and investigated.

2. Contributions

The special issue attracted 19 contributions in total. After an extensive review period involving the distinguished expert

reviewers and the guest editors, five papers were accepted for the special issue, resulting in an acceptance rate of 26%. The introduction of the accepted papers is presented below.

The first paper, titled "Analysis of Denial of Service Impact on Data Routing in Mobile eHealth Wireless Mesh Network," by S. Alanazi et al., states that wireless mesh networks (WMNs) are a promising technology that has emerged with the combination of several wireless networks. These wireless networks and devices communicate in a mesh network manner, to provide edge-to- edge, easy, and cost-effective data communication. Many current and future promising applications depend on WMNs and one of the most important is eHealth, where the confidential information is transmitted with the help of WMNs. The authors state that denial of service (DoS) attacks are fatal to many types of networks, including wireless mesh networks, specifically when the network is utilized in a highly sensitive scenario like eHealthcare. This paper analyzes three types of attacks that can cause DoS in static and mobile WMNs and the remedies against them.

The second paper, titled "A Case of Engineering Quality for Mobile Healthcare Applications Using Augmented Personal Software Process Improvement," by S. A. K. Ghayyur et al., starts with the discussion that mobile healthcare systems are currently considered among the key research areas in the domain of software engineering. Modern technologies, for mobile healthcare systems, are readily available. The authors present the Architecture Augmented Personal Process technique in order to enhance the quality of mobile healthcare systems through the use of an architectural design with an integration of the personal software process.

The third paper, titled "Integrated Wearable System for Monitoring Heart Rate and Step during Physical Activity," by E. A. P. J. Prawiro et al., proposes integrating a heart rate (HR) monitoring system with a step counter for use during physical activities. A novel step counter algorithm has been developed to enable the highly accurate detection of a step. The proposed system comprises a wireless wearable device, a smartphone, and a remote server. Data transmission between a wearable device and a smartphone is conducted via Bluetooth low energy (BLE). An indirect contact measurement method has also been devised to eliminate the need for direct contact electrodes and the likelihood of skin irritation. The proposed system is compact, lightweight, and comfortable to wear. A smartphone application provides the interface for the display of data related to HR, step count (SC), exercise intensity, speed, distance, and calories burned, as well as waveforms related to ECG and step cycle. The ECG peak detection algorithm achieved an accuracy of 99.7% using the MIT-BIH ST change database. An accuracy of 98.89% was achieved for HR and 98.96% for SC at treadmill speeds of 1.8 to 9.0 km/h.

The fourth paper, titled "Automatic Gender Detection based on Characteristics of Vocal Folds for Mobile Healthcare System," by M. Alhussein et al., proposes that automatic gender detection may be useful in some applications of a mobile healthcare system. In a human voice production system, the contribution of the vocal folds is very vital. The length of the vocal folds is gender dependent; a male speaker has longer vocal folds than a female speaker. Due to longer vocal folds, the voice of a male speaker becomes heavy and, therefore, contains more voice intensity. Based on this idea, a new type of a time domain acoustic feature for an automatic gender detection system is proposed in this paper. The proposed feature measures the voice intensity by calculating the area under the modified voice contour to make the differentiation between males and females. Two different databases are used to show that the proposed feature is independent of text, spoken language, dialect region, recording system, and environment. The experimental results show that the detection rates for clean and noisy speech are 98.27% and 96.55%, respectively.

The fifth and the last paper, titled "EVFDT: An Enhanced Very Fast Decision Tree Algorithm for Detecting Distributed Denial of Service Attack in Cloud-Assisted Wireless Body Area Network," by R. Latif et al., states that the detection denial of service attacks demands an adaptive and incremental learning classifier capable of accurate decision making with less computation. The DDoS attack detection using existing machine learning techniques requires the full data set to be stored in the memory and is not appropriate for real-time network traffic. To overcome these shortcomings, the Very Fast Decision Tree (VFDT) algorithm has been proposed in the past that can handle high speed streaming data efficiently. While considering the data generated by WBAN sensors, noise is an obvious aspect that severely affects the accuracy and increases false alarms. In this paper, an enhanced VFDT (EVFDT) algorithm is proposed to efficiently detect the occurrence of DDoS attacks in cloud-assisted WBANs. EVFDT uses an adaptive tie-breaking threshold for node splitting. To resolve the tree size expansion under extreme noise, a lightweight iterative pruning technique is proposed. To analyze the performance of EVFDT, four metrics are used: classification accuracy, tree size, time, and memory. Simulation results show that EVFDT attains a significantly high detection accuracy with fewer false alarms.

Acknowledgments

We are grateful to the authors who submitted their valuable contributions for consideration in this special issue, and the reviewers who generously donated their time in the review process. Without their help and contributions, this special issue would not have been possible. We hope that the readers will find the papers in this special issue stimulating.

Basit Shahzad
Mehmet A. Orgun
Christoph Thuemmler

Hindawi Publishing Corporation
Mobile Information Systems
Volume 2016, Article ID 4853924, 19 pages
http://dx.doi.org/10.1155/2016/4853924

Research Article

Analysis of Denial of Service Impact on Data Routing in Mobile eHealth Wireless Mesh Network

Shaker Alanazi,[1,2] Kashif Saleem,[1] Jalal Al-Muhtadi,[2] and Abdelouahid Derhab[1]

[1]*Center of Excellence in Information Assurance (CoEIA), King Saud University, Riyadh 12372, Saudi Arabia*
[2]*College of Computer & Information Sciences, King Saud University, Riyadh 12372, Saudi Arabia*

Correspondence should be addressed to Kashif Saleem; ksaleem@ksu.edu.sa

Received 12 January 2016; Accepted 8 June 2016

Academic Editor: Laurence T. Yang

Wireless mesh networks (WMNs) are a promising technology that has emerged with the combination of several wireless networks. These wireless networks and devices communicate in a mesh network manner, to provide edge-to-edge, easy, and cost-effective data communication. Many current and future promising applications depend on WMN and one of the most important applications is eHealthcare, where the confidential information transfers with the help of WMN. WMN devices communicate over a wireless medium, which opens the system to a number of vulnerabilities; thus, an intruder can launch malicious activities through many types of attacks that can result in denial of service (DoS). In this paper, the available solutions to overcome these attacks are simulated and evaluated in terms of data packet delivery ratio, end-to-end delay, and network throughput and under different cases of static and mobile WMNs, which helps in providing suggestions to enhance existing protocols and mitigate the effect of DoS caused by such attacks.

1. Introduction

Security is always a major concern in transferring information from the source to the destination, whether it is the manual postal service or data communication in today's digitized world, especially, applications like eHealthcare that are built to provide assistance for a very sensitive purpose. Over the years, many communication technologies have been introduced and numerous protocols have been proposed to handle security issues. One promising and growing wireless communication technology is the wireless mesh network (WMN). In eHealthcare, WMN plays a very important role in transferring confidential information hop by hop until the required destination. A WMN is a self-organized and self-configured network, where nodes and networks communicate in a mesh network manner to provide connectivity [1]. A WMN consists of nodes that communicate using wireless technology to provide connectivity. These nodes are usually mesh clients such as cell phones and mesh routers, including Wi-Fi access points. A WMN has a decentralized architecture where mesh clients and routers communicate in a distributed manner to provide the connection [2]. The authors of [3] discussed the main design factors that affect the architecture of a WMN and the selected components to build the WMN, such as signal transmission techniques, scalability, connectivity, the quality of broadband services, security, ease of use, and compatibility. WMN nodes can be connected and organized in three main ways to form three types of WMNs: infrastructure/backbone, client, and hybrid WMNs. The main characteristics and benefits of WMNs are coverage, mobility, scalability, heterogeneity, and compatibility [3] and provide network services such as Internet access, video conferencing, and voice communication, and they help in transferring data between networks. This makes a WMN suitable for many applications in both military and civilian fields. Public safety and disaster recovery (PSDR) wireless communication systems are an example where a WMN is strongly applicable [3, 4]. In such applications, wired communication may not be possible or risky, so a WMN is a more suitable solution because it offers easy and rapid connectivity through a wireless medium.

A WMN can be seen as a group of nodes (clients or routers) that cooperate to provide connectivity. Such an open architecture, where clients serve as routers to forward

data packets, is exposed to many types of attacks that can interrupt the whole network and cause denial of service (DoS). Moreover, the mobility factor in the network makes the security of the network more difficult and challenging. For this reason, protection against DoS problems is essential and requires secure routing protocols. These secure routing protocols must detect, identify, and isolate attacks that can cause DoS, and they must keep the routing functionality working.

Very common and serious attacks on WMN routing that can lead to DoS are selective forwarding, hello flooding, and wormhole attacks [17, 18]. In a selective forwarding attack, the packets that pass through malicious nodes are routed by these nodes inconsistently to distribute the network. In a hello flooding attack, a malicious node transmits hello messages with high transmission power to convince the remaining nodes that it is a neighbor to many nodes in the network. A wormhole attack is one of the most dangerous attacks in a WMN network layer. In this type of attack, a malicious node tunnels messages from one part of the network to another part and replays them. To make the situation worse, a high-speed link with low latency is usually used to tunnel these messages, which makes the link more tempting to be selected in the routing protocols. All of these types of attacks aim to disrupt the routing process to cause the DoS problem. A selective forwarding attack tries to cause DoS directly by dropping some of the routed packets. In contrast, in hello flooding and wormhole attacks malicious nodes try to give themselves a greater chance of being selected by the routing protocol as a routing path; then they can perform any activity to distribute the network and introduce DoS.

Even though the WMN is a promising technology, it faces many challenges, some of which are mentioned above. This paper studies the impact of three recent and common types of attacks on mobile WMNs: selective forwarding, wormhole, and hello flooding attacks. Moreover, this paper will study existing solutions to overcome these attacks on the mobile environment. A part of this work was presented previously in [19] as a preliminary report.

To achieve the objective, the three types of attacks are simulated using OMNET++ on Ad hoc On-Demand Distance Vector Routing (AODV). AODV was selected because it is proposed as a possible routing protocol in IEEE 802.11s [20]. Moreover, AODV is available in many simulation tools, including OMNET++. The attacks are simulated on the OADV routing protocol in two modes. In the first mode, malicious nodes are stationary, and, in the second mode, the network contains mobile malicious nodes. In both modes, the impact of the attacks on the packet delivery ratio, end-to-end delay, network throughput, and energy consumption are measured. This will highlight the impact of these attacks on the network, as well as the effect of mobility on increasing or decreasing the damage caused by the attacks.

After studying the impact of the attacks on stationary and mobile networks, the study will concentrate on the available solutions discussed in the literature to overcome the attacks. More specifically, the same simulations and experiments performed on AODV will be conducted on a security protocol called Position-Aware, Secure, and Efficient Mesh Routing (PASER). This will measure PASER's immunity to the three types of attacks on both stationary and mobile networks. The research will contribute to the field as follows:

(i) Implementing 28 scenarios that include AODV and the WMN protocol PASER in OMNET++ for both malicious and legitimate scenarios.

(ii) Evaluating the AODV and PASER protocols by incorporating selective forwarding, wormhole, and hello flooding attacks in a mobile WMN; to the best of our knowledge, this study is the first of its kind.

(iii) Evaluating and measuring the effectiveness of the two principle protection mechanisms, cryptography and a GPS module, against attacks in a mobile network.

(iv) Providing suggestions that facilitate future research studies to simulate different WMN scenarios and the adoption of appropriate security mechanisms.

This paper is structured as follows. Section 2 discusses the issues and challenges faced by WMNs, reviews the related and recent literature, and gives comparison of the available solutions. Section 3 presents implementation and results of attacks on AODV and PASER. Section 4 exhibits the discussion and recommendations. The conclusion and future work are provided in Section 5.

2. Literature Review

The routing layer in a WMN, where packets travel from the source to the destination in a multihop manner, is one of the main differentiators of a WMN from other types of networks. Such a mechanism makes a WMN more vulnerable to network attacks, such as black holes, wormholes, grey holes, and packet dropping. Although many secure routing protocols have been proposed to address the security vulnerabilities, there is still significant room to enhance the security and usability of WMN secure routing protocols. In this section, the current state-of-the-art secure routing protocols are reviewed.

2.1. Issues and Challenges Faced by WMNs. A wireless mesh network (WMN) is a type of network where nodes cooperate and transfer data packets hop by hop. In hop-by-hop communication, network security is the responsibility of not only the routers but also each node in the network. When regular nodes contribute to communication over the wireless medium, they are vulnerable to many types of attacks that can lead to a denial of services (DoS) in the WMN. Moreover, when mobile nodes are joining or leaving the WMN, the problem becomes more complicated because a trust issue can arise between the mobile nodes and the network.

Still, there are open challenges and issues in WMNs to be addressed by the research community and industry, as these challenges affect the usability of WMN technology in most applications. In [1], the authors discuss the open issues and challenges of WMNs at every layer. One of these open issues in the routing layer is the need for a protocol that is efficient specifically for WMNs. As described in [1], efficient

routing protocols are distributed and independent of any traffic profile, and they feature link quality variation and minimum overhead.

Some of the most important challenges faced by WMNs are network provisioning and network integration. In network provisioning, there is a need for a sophisticated management tool to enable mesh routers and mesh clients to dynamically establish connections and to manage the mobile nodes. Since a WMN contains various types of technologies and protocols, it is necessary to have standard mechanisms to integrate these technologies. For example, some mesh routers are based on IEEE 802.11, while others are based on different standards, such as IEEE 802.15.4 and IEEE 802.16.

Security is one of the most important issues that need to be addressed in WMNs. Unfortunately, WMNs are vulnerable to multiple types of attacks because of their special characteristics, such as their ad hoc nature and use of a wireless medium to communicate [4].

In the physical layer, WMNs are vulnerable to jamming attacks, according to [17], for example, trivial signal jamming where the attacker transmits a noise signal to disrupt the WMN radio signal. In addition, a WMN's physical layer is vulnerable to reactive jamming, where the attacker transmits noise whenever it detects that a legitimate node has started to use the channel. Moreover, in some applications of WMNs, the installation or usage environment is not secure, such as in remote areas, which gives attackers the ability to tamper with the system to extract important information or even destroy the units.

The MAC layer in a WMN is vulnerable to many types of attacks. Some examples provided in [17] are jamming attacks and MAC spoofing. In jamming attacks, an attacker can try to disrupt the communication channel by sending complete MAC frames. An unprompted Clear-to-Send (CTS) attack, a reactive Request to Send (RTS), and corrupt jamming (CJ) are some examples of possible MAC jamming attacks in a WMN's MAC layer.

The network layer of a WMN is also vulnerable to multiple types of attacks that may lead to the DoS problem. In [17], examples of attacks are given that can target the network layer in the control or data plan. In the control plan, rushing, routing table overflow, Sybil, wormhole, and sinkhole attacks are some examples. The most basic data plan attack is eavesdropping, where a malicious node tries to learn the network topology by listing to the traffic.

Many types of attacks can be launched in the transport and application layers of a WMN. In the transport layer, an attacker can try to flood the network by repeatedly issuing connection requests to a specific resource in the network [17]. The application layer is also vulnerable to many types of attacks, including viruses, worms, and malicious codes and application abuse [17].

A great deal of effort has been put into treating WMN security issues and proposing suitable solutions at many levels and in various applications. For example, in [18, 21], the authors study the security requirements and the impact of implementing security measures in WMNs in general, while, in [22], the performance of the routing protocol of an 802.11-based WMN under attack is studied. Many solutions have been proposed to handle WMN security requirements, either to solve a specific challenge, such as the key establishment mechanism proposed in [23], or by providing a more comprehensive solution, as discussed in the next section. The next section discusses recent security solutions and secure routing protocols developed to protect the WMN network layer, mainly against attacks that can lead to DoS.

2.2. Recent Related Work. The authors in [5] propose a routing protocol to protect WMNs against wormhole attacks called "wormhole-resistant secure routing for wireless mesh network (WRSR)." The authors of the paper followed a new approach that aims to detect the presence of wormhole nodes and links and quarantine them before using the links. WRSR depends on the neighbors' information and the existence of alternative subpaths to detect the existence of a wormhole. WRSR uses statistical probability to categorize the paths as either safe, "wormhole free," or unsafe, "containing a wormhole." In the proposed protocol, WRSR represents the transmission range of a node and d is the distance between two nodes. The authors use a unit disk graph to prove that the probability of finding an alternative path where $R < d < 2R$ is high, and any path that does not fit in this category is considered insecure. To implement the protocol, IEEE 802.11s frames are extended to contain the necessary information to make WRSR work, which are a flag bit to indicate the existence of neighboring information and the neighboring addresses to accommodate various neighbors' numbers. To simulate and test the proposed protocol, a strong adversary model is used where malicious nodes can establish links with high speed and low latency. Moreover, the attacker can compromise mesh routers in the network to lunch attacks. In the simulation, the protocol shows a very high correct rate of detection of wormhole links when the network is dense and contains many alternative subpaths.

In [6] the authors have proposed an efficient and secure routing protocol for a hybrid wireless mesh network (WSRPHWM) that depends heavily on public cryptography to secure routing. First, it is important to mention that WSRPHWM is based on a route-on-demand protocol called the cross-layer secure and resource-aware on-demand routing protocol (CSROR), where WSRPHWM inherits the routing metrics used in CSROR. The routing metrics used depend on three factors to quantify routes: the available bandwidth, node battery power, and unreliability level. The unreliability level is based on value assigned by a node's neighbor to indicate the past experience of forwarding packets on a particular route. To maintain security, WSRPHWM uses symmetric key cryptography, asymmetric key cryptography, and the MAC function to authenticate and encrypt routing data. WSRPHWM assumes that the network has a certification authority (CA) that issues and signs certificates for each node and router in the network. The first step taken by each node is to establish a secure session using the Diffie-Hellman elliptic curve with each of its neighbors to generate a trusted session key between them. After that, for each message that includes a route request, the sender node encrypts the mutable fields of the request using the session keys for each neighbor, signs the nonmutable fields using its private key, and finally generates

a MAC using the mutual session key. For each neighbor, the process is repeated. When the neighbor receives the request, it first validates the MAC and the signature of the packet and then decrypts the mutable fields to alter them. After that, it repeats the process done by the source node and then forwards the request to its neighbors. When WSRPHWM is compared with other protocols, such as SADOV and CSROR, it shows very good results with respect to the average end-to-end delay and routing overhead. For the average end-to-end delay, WSRPHWM shows better performance than SADOV, but the delay is longer than that of CSROR because it uses the crypto operations. WSRPHWM causes the smallest routing overhead among the three protocols.

The cross-layer secure and resources-aware on-demand routing protocol (CSROR) for hybrid wireless mesh networks proposed in [7] uses the history of communication to decide which route is trustworthy. To do so, the route selection parameters include a threat level (TL) field in addition to the power level and available bandwidth. The TL of the route is the sum of the drop values (DV) of nodes in the route, while the DV of each node is the percentage of unacknowledged messages. In CSROR, each node keeps the drop value of each of its close neighbors (one level). The DV is calculated by subtracting the number of packets passively acknowledged by the direct neighbor of the current node neighbor. When the route discovery process starts in CSROR, the source node broadcast route request packet (RREQ) and the destination select the best route according to the aforementioned parameters. CSROR shows better results than AODV and ASODV in terms of efficiency and reliability, particularly in the packet delivery rate and average delay.

In [8], the authors propose a routing scheme based on the ADOV distance vector algorithm to provide secure routing, using different control packets for routing requests and routing request replies. The selection of the route in the proposed routing protocol is based on two factors: the data rate and node reputation value (RV). The selection of the route is done by the destination when it receives routing request packets. The route that has the highest reputation and the lowest data rate is selected. The RV of each node is generated by its neighbors to signify previous successful communication, and the available data rate is calculated by exchanging information between the network and the MAC layer.

Authenticated Routing for Ad hoc Networks (ARAN) in [9] is another routing protocol that depends heavily on cryptography to provide security. ARAN requires the existence of a certificate server responsible for issuing, distributing, and revoking certificates. In ARAN, each node must have a certificate signed by and trusted by this server to be considered trusted. The certificate of each node must contain the IP address, its public key, a time stamp to indicate when the certificate was created, and the expiry date. In the route discovery process, the source signs a route request packet (RREQ) and broadcasts it to its neighbors. The neighbors first validate the packets and then sign and broadcast them. When the destination receives the request packet (RREQ), it sends a reply using the reverse path. Moreover, it signs the reply and includes its certificate in the reply, and the same sequence of signing and verifying that took place in the route discovery is repeated until the reply reaches the source. Finally, when a certificate is to be revoked, the trusted server will broadcast a signed request for the certificate to be revoked. Each node rebroadcasts the request to its neighbors.

The SAODV protocol proposed in [10] aims to secure AODV using cryptography. In fact, it solves many security flaws in AODV. The first flaw is the possibility for an attacker to impersonate an S node by forging route request packets (RREQ) using its address as the source address. The second security flaw is that the hop counts in RREQ packets are reduced to select the best route. The third flaw is the ability to impersonate a destination node by replying to RREQ as if it is the destination. To solve these security issues, SAODV uses cryptography, mainly hashing and digital signatures. SAODV assumes the existence of a key management system that is responsible for issuing nodes' keys and revoking them if needed. It is important here to mention that these keys or certificates must associate the key with the address of the nodes, which is a challenge because of the dynamic nature of WMN networks. Once all nodes have their public keys or certificates, they can start communication. In ADOV, the RREQ hop count is the only mutable field that can be secured using a hashing chain. The nonmutable fields in RREQ and RREP can be secured using a digital signature.

A modified version of Secure Efficient Ad hoc Distance (I-SEAD) vector routing is proposed in [11] based on the destination sequenced distance vector (DSDV) with the enhancement of adding security to the routing. I-SEAD uses a one-way hash chain to authenticate the routing update. To build the one-way hash chain, I-SEAD uses a cryptographic hash function; the one-way hash chain is computed by conducting consecutive hashing, $h_i = H(h_{i-1})$ for $0 < i \leq n$, for some n and when the first value in the chain, x, is randomly generated for each node. This process enables nodes to determine whether the updates came from a legitimate node. I-SEAD assumes the existence of trusted distribution authority to distribute the initial values of nodes.

In [12], the Security-Aware Ad hoc Routing (SAR) for wireless networks is discussed. SAR includes security as a parameter in the route discovery process. Moreover, it provides a security framework that can be adapted by any application that uses SAR according to the required level of security. SAR protects the route discovery process by allowing only nodes that have at least the same level of security to contribute to the process. The required level of trust is embedded in the route request packet (RREQ); nodes that cannot satisfy the required level of security cannot forward or reply to the route request. Moreover, SAR uses cryptography to provide message authentication, message integrity, message confidentiality, and node authentication.

In [13], the designers propose a new methodology to detect malicious routing activity by colluding nodes in WMNs. The new methodology is not an independent and complete secure routing solution; it is a mechanism to be integrated into routing protocols to detect malicious routing activities. The proposed solution, called Leak Detector, is based on graph theory, where the destination node of the route builds a virtual graph to model the paths from the

source to the destination. Using this virtual graph, the destination node calculates the ratio of the incoming and outgoing traffic for each intermediate node in the path and uses this information to decide whether an intermediate node is malicious or not. When the deviation between the incoming traffic and outgoing traffic is high, this node is considered malicious. The proposed mechanism detects only selective forwarding or the complete dropping of packets, and it assumes the existence of multiple paths from the source to the destination.

A secure routing protocol against wormhole attacks in a sensor network (SeRWA) is described in [14]. SeRWA uses neighbor information to detect wormholes. In terms of security, SeRWA provides a way to detect wormholes and provides authentication between neighbors using a pair-wise key. To detect wormhole tunnel, each node builds a list of its neighbors and shares this list with all its neighbors. When a neighboring node exists in the node's routing table and does not exist in its routing table of neighbors, the protocol uses the distance between these neighbor nodes and transmission power to decide whether the missing node is part of a wormhole tunnel. A simulation of SeRWA is performed using NS2 and shows a low rate of false positive detection of less than 10%.

In [15], the authors propose a secure routing protocol against routing disruption in MANET networks called CRP. The proposed protocol is based on dynamic source routing (DSR) and uses public key cryptography to detect tampering with packets or fake packets. The proposed protocol assumes the existence of monitoring nodes to monitor intermediate nodes between monitoring nodes and collect statistics about the routing behavior of these nodes to detect black hole and grey hole attacks. The proposed protocol is simulated using NS2 and compared to the DSR protocol. For a black hole attack in which malicious nodes form about 40% of the network, CRP achieves a more than 70% packet delivery rate, while DSR achieves less than 50%. For a grey hole attack, CRP achieves about an 80% packet delivery rate and DSR achieves about 70% when 30% of the network nodes are malicious.

2.2.1. Position-Aware, Secure, and Efficient Mesh Routing (PASER). In [16], the authors propose a routing protocol called Position-Aware, Secure, and Efficient Routing (PASER) discovery protocol for wireless mesh networks. PASER is a reactive routing protocol that uses cryptographic operations to protect the routing of packets in mesh networks. The designers of PASER stated three main goals to be accomplished by PASER.

(i) *Protection against External Attacks.* PASER uses public key cryptography to identify nodes and give the nodes the ability to take part in the routing mechanism. This is accomplished by the existence of a key distribution center (KDC) that issues certificates for trusted nodes. This can be compromised only by compromising the session keys or node certificates.

(ii) *Isolate and Exclude Malicious Nodes from the Network.* Even though this is one of objectives of PASER, the designers did not implement a specific mechanism to accomplish this; instead, this is done by the route discovery process. The designers of PASER discussed a variety of options that can be implemented, including the adoption of a honeypot.

(iii) *Reducing the Impact of Malicious Nodes in the Network.* This objective is accomplished by using cryptographic functions to limit the operations that can be performed by malicious nodes in the routing process. For example, message authentication is used to prevent malicious nodes from faking or altering messages. Message freshness is also used to prevent the replaying of old messages. Moreover, origin authentication is implemented by PASER to verify the origin of a message. One important protection mechanism in the routing process is neighbor authentication to ensure that the neighbor is in transmission range, which is the main defense mechanism against wormhole attacks. This is done in PASER by associating the digital signature of the neighbor with the neighbor's GPS location.

There is an existing implementation of the complete protocol that can be integrated to OMNET++, where OpenSSL is used as the encryption library.

2.2.2. Ad Hoc On-Demand Distance Vector Routing (AODV). According to AODV RFC 3561 in [24], AODV is a reactive routing protocol that establishes the route only when there is data to be sent. AODV has three main advantages that make it attractive to be used in WMNs:

(i) It is loop-free.

(ii) AODV requires little bandwidth because the routing table and control data are very small.

(iii) It is very scalable.

In general, AODV uses four types of routing messages.

(i) *Route Request (RREQ).* This message is broadcasted by the source requesting the route and forwarded by the intermediate nodes if the message has not been processed before.

(ii) *Route Reply (RREP).* A message is sent to identify that a route has been found to the destination. This message is sent by a node; if it receives the RREQ message, it is the destination node or in the path of the destination node, and the message has a sequence number smaller than or equal to the one in the RREQ message.

(iii) *Route Error (RERR).* This message is sent by one of the nodes in the path used, source node, or destination node. The message is sent when the link breaks to notify the source to initiate the route discovery process again by broadcasting an RREQ message.

(iv) *Route Reply Acknowledgment (RREP-ACK).* This is used by the sender to ensure the availability of a link from the destination to the sender.

2.3. Comparison. The aforementioned reviews of secure routing protocols for WMN are summarized in Table 1, which highlights the core differences and limitations. These protocols vary in many respects, such as the mechanism architecture, tackled attacks, the technique used to detect the attacks, and efficiency. Four parameters are considered as limitations in the comparison below, which are complexity, requirement of additional resources, overhead, and efficiency of the protocol in protecting the network against attacks which might lead to DoS.

The comparison shows us that most secure routing protocols utilize cryptography [6, 9–12, 14, 16] to provide data security such as authentication, data confidentiality, and integrity over the network. Unfortunately, these protocols do not guarantee the protection of WMNs against the most important attacks, for example, selective forwarding, wormhole, and hello flooding [9, 11, 12], which jams data traffic and thus causes denial of service (DoS). Furthermore, some protocols [6, 8–10] assume that an authority already exists in the network to manage the distribution of certificates for data encryption.

Aside from cryptography and CA, two other types of solutions [5, 7, 8] have been proposed to handle the DoS problem: hardware-based and statistical-based solutions. In hardware-based solutions, GPS or an antenna is used to verify the location of a node and validate that it is legitimate neighbor [9]. In statistical-based solutions, the nodes keep track of their successful past communication with neighbors to select the best next hop [5, 7]. In general, solutions against attacks that target WMNs are still an open issue [2, 4, 17, 18].

3. Implementation and Results

The selection of the routing protocol to be used in the current research is critical since the study focuses on attacks on the routing protocol of WMNs.

3.1. Routing Protocols and Attacks. The selected routing protocol to be used in this study is AODV. AODV implementation is available for simulation tools, including OMNET++, which was selected to be used in simulation of these attacks because of its suitability [25].

In OMNET++, the INET framework facilitates the simulation and contains modules for wired and wireless communication, including the AODV routing protocol. INETMANET is an extension of the INET package that contains more routing protocols and standards related to mobility and wireless networks, including WMNs.

The second selected routing protocol is Position-Aware, Secure, and Efficient Mesh Routing (PASER). The reasons for selecting PASER for the current research are as follows.

(i) *Availability of Resources*. The PASER design and development team has put a great deal of effort into making the protocol available for researchers as well as for industry. This is reflected by the implementation of the protocol in C/C++ and integration for use in a Linux environment as a complete implementation where it can be ported and tested on real hardware or used in OMNET++ as a simulation package. For the simulation, it is available for OMNET++ for both Linux and Windows operating systems and integrated with the INET package. The only missing part of the resources is the availability of full documentation to describe in detail the PASER structure and how to use the protocol with OMNET++ packages.

(ii) *Usable Components*. The PASER design team used multiple components in their designs that are parts of other routing protocols or well-known packages, which makes the results of the study more accurate and trustworthy. In regard to routing methodology, PASER adapted the route discovery process from the AODV routing protocol. In addition, the route maintenance and detection process in PASER is adapted from the neighborhood discovery protocol (NHPS). Cryptographic primitives are a building block in PASER that were adapted from a well-known package in industry and academia, OpenSSL.

3.1.1. Hello Flooding Attacks. A hello flooding attack [26, 27] is a very destructive DoS attack. The malicious node tries to give itself a greater probability to be selected as a route by making itself more appealing. In wireless networks, this is usually accomplished by transmitting using a high power frequency to its neighbors, which causes the malicious node to cover a larger area and provide a shorter path to the destination [28]. After being selected as a node in the route, the malicious node then performs malicious activities on the received packets, such as altering packets or gain statistics or simply dropping the packets [29]. In the current research, the implementation of a hello flooding attack is performed as follows.

(i) *High Transmission Power*. The malicious node is configured to have higher transmission power than legitimate nodes. This is done by changing the .wlan* .radio.transmitterPower attribute in OMNET++ initialization file.

(ii) *Dropping the Packets*. The packets that go through the malicious node are simply dropped; this helps to gain more insightful information to study the effect of the attack. To do this correctly, the malicious node must drop only data packets, not routing and control packets; otherwise, the malicious node will not take part in the routing discovery process and will not be part of the route. Thus, the source code for the routing protocol in the malicious node must be altered by dropping the packets rather than forwarding them as shown in Algorithm 1.

3.1.2. Selective Forwarding Attack. A selective forwarding attack or grey hole attack is less destructive than a hello flooding attack because the malicious node has no advantage over legitimate nodes to be selected in the route. In a selective forwarding attack, the packets are forwarded inconsistently according to the attacker's goal [30], which can cause DoS in the network [31–33]. In the current research, the malicious node forwards the packets with a probability of 50%; otherwise, the packets will be dropped as shown in Algorithm 2.

Table 1: Comparison of the available solutions.

Reference	Mechanism	Tackled attacks	Security analysis	Limitations
WRSR [5]	Statistical based: neighbors' information and existing of alternative path	Wormhole	Simulation with strong adversary scenario	Inefficient protection due to limited number of attacks handled
WSRPHWM [6]	Cryptography based and past communication statistics	Spoofed route signaling Replay attack Black hole Wormhole Grey hole Routing disruption	Theoretical	(i) Additional resources are required which is a certification authority (ii) Overhead due to usage of public key cryptography
CSROR [7]	Past communication statistics	Wormhole, black hole, and grey hole	Not provided	Inefficient protection due to limited number of attacks handled
E-SRPM [8]	Link's length information and random walk route scheme	Wormhole	Simulation	Inefficient protection due to limited number of attacks handled
ARANA [9]	Cryptography based	Malicious packets manipulation	Simulation	(i) Additional resources are required which is a certification authority (ii) Overhead due to usage of public key cryptography (iii) Inefficient protection due to limited number of attacks handled
SAODV [10]	Cryptography based	Malicious packets manipulation of routing metric and nodes impersonation	Simulation	(i) Additional resources are required which is a certification authority (ii) Overhead due to usage of public key cryptography (iii) Inefficient protection due to limited number of attacks handled
I-SEAD [11]	Cryptography based	Malicious routing update	Simulation	Inefficient protection due to limited number of attacks handled
SAR [12]	Cryptography based	Route discovery process	Theoretical	(i) Additional resources are required which is a certification authority (ii) Overhead due to usage of public key cryptography (iii) Inefficient protection due to limited number of attacks handled
Leak detector [13]	Statistical based: neighbors' information and existence of alternative path	Selective forwarding and black hole	Simulation	(i) Inefficient protection due to limited number of attacks handled (ii) Complexity due to required integration in existing routing protocol
SeRWA [14]	Neighbor information and cryptography based	Wormhole	Simulation	(i) Additional resources are required to distribute keys (ii) Inefficient protection due to limited number of attacks handled
CRP [15]	Statistical based: neighbor information and public key cryptography	Black hole, grey hole, packet tampering, rushing attack, and collusion attack	Simulation	(i) Additional resources are required which is a certification authority (ii) Overhead due to usage of public key cryptography (iii) Inefficient protection due to limited number of attacks handled
PASER [16]	Cryptography based and GPS location	Wormhole and cryptographic protection (message and node authentication, message freshness, and message confidentiality)	Simulation	(i) Additional resources are required which is a certification authority and GPS hardware (ii) Overhead due to usage of public key cryptography

```
If (isHelloFlooding)
{
   if (dynamic_cast<UDPPacket *>(msg))
          {
             delete msg;
             return;
          }
}
```

Algorithm 1: Hello flooding attack implementation.

```
if (MIPs.equals(getAddress( ).getIPv4( )))
{
   if (uniform(0,1) < 0.5 & is Selective)
   {
      if (dynamic_cast<UDPPacket *>(msg))
          {
             delete msg;
             return;
          }
   }
}
```

Algorithm 2: Selective forwarding attack implementation.

3.1.3. Wormhole Attack. A wormhole is a more sophisticated attack that requires at least two malicious nodes to perform the attack by building wormhole tunnel [34]. This tunnel is usually built in wired networks by connecting a wire from the source to the destination; in the case of wireless communications the tunnel is built by having high-frequency overlapping coverage ranges for both the source and the destination [35, 36]. The source tries to be more attractive as a path selection for the packets by convincing its neighbors that it sees part of the network they cannot see which is a shorter path to the legitimate destination as shown in Figure 1. In the current research, the tunnel is built by high-frequency transmission power for both the tunnel source and destination.

3.2. Network Design and Parameters. This section describes the network parameters and configuration for the experiment. Table 2 shows the general configuration of the network, while Table 3 shows the specific network configuration related to the attack simulations. Tables 4 and 5 show the initial battery configuration and battery consumption parameters, respectively.

Figure 2 is the diagram that shows the initial layout and setup after configuring the above parameters.

The coverage of the radio signal in each node is a critical parameter in the research for the following reasons.

(i) It is used in routing protocol to find a path to the destination.

Table 2: General network configuration.

General configuration	
Simulation time	70,000 s
Number of nodes	50
Max area (x-axis)	1,000 M
Max area (y-axis)	1,000 M
Min area (x-axis)	0 M
Min area (y-axis)	0 M
Network layout for intermediate nodes	Random
Position for the source x-axis	1 M
Position for the source y-axis	1 M
Position for the destination x-axis	999 M
Position for the source y-axis	999 M

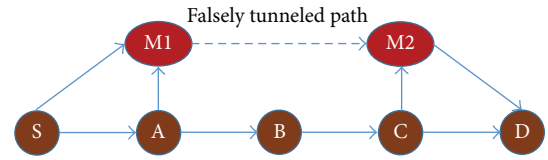


Figure 1: Wormhole attack description.

(ii) It is used by malicious nodes to attract other nodes to forward the traffic through them (in wormhole and hello flooding attacks).

Figure 3 shows the initial coverage of each node.

One of the goals of this research is to study the effect of mobility when malicious nodes are generating attacks in the network. For this reason, the mobility configuration plays a crucial part. The parameters of mobility for the two malicious nodes, which are numbered 1 and 2, are listed in Table 6.

3.3. Adversary Model. The adversary model and attacks scenarios as shown in Table 7 are performed on AODV and PASER protocols.

3.4. Simulation Scenarios with Malicious Nodes Performing a Hello Flooding Attack. In this section, the effect of a hello flooding attack on stationary and mobile networks is measured. The attack is performed by the transmission of high-frequency hello messages. Packet delivery ratio (PDR), throughput, and end-to-end delay are calculated and graphs are generated by Omnet++. After conducting attacks as described in Section 3.3, the collected results are shown in Table 8.

As Figure 4 shows, network PDR is strongly affected by hello flooding attack; the highest drop in PDR happened when attack scenario number 8 was conducted. This scenario caused PDR to decrease from 94.4% when the network does not contain any malicious node to reach 71.15% when eight mobile malicious nodes were conducting the attack. In summary, it can be said that the network PDR is strongly affected by hello flooding attack and the PDR decreases when the number of malicious nodes increases. Moreover, mobile

Table 3: Routing protocol parameters.

Network parameters	
Communication type	Wireless
MAC Protocol	IEEE 802.11g
Type of traffic	UDP
Routing protocol	AODV, PASER
Type of IP	IPv4
Source port	1234
Distention port	1234
Packet length	512 Bytes
IP forwarding	Enabled
IGMP type	IGMPv2
Packet time to live	32
Sending intervals	Random
Sleeping duration	1 second
Interface start time	10 seconds
Number of radio interface	1
Wireless NIC bitrate	54 Mbps
Wireless NIC frame capacity	10
Maximum queue size	14
Basic bitrate	6 Mbps
Sending retry limit	7
Radio maximum transmission power	20 mW
Maximum transmission power	6.0 mW
Default radio transmission power	2.0 mW
Radio sensitivity	−90 dBm
Broadcast delay	Random between 0 s and 0.005 s
Radio carrier frequency	2.4 GHz
Propagation model	Free space model

Table 4: Initial battery configuration.

Battery configuration	
Battery type	InetSimpleBattery
Battery nominal	25
Battery capacity	25
Battery voltage	10
Battery resolution	1 s
Delta value of battery	0.5
Battery publishing time	20 s
Battery consumption factors for radio and CPU	(i) Module is idle (ii) Module is asleep (iii) Module is sending packets (iv) Module is receiving packets (v) CPU is active (vi) CPU is standing by

Table 5: Battery consumption parameters.

Action	Cost
Radio, idle	1.3 mA
Radio, receive	9.0 mA
Radio, sleep	0.06 mA
Radio, send	9.0 mA

Table 6: Mobility configuration.

Node number	Mobility configuration	
Node number 1	Mobility Type	Circle mobility
	Radius	150 m
	Speed	100 mps
	Start Angle	120 degrees
	mobility.cy	250 m
	mobility.cx	250 m
Node number 2	Mobility Type	Constant speed mobility
	Speed	50 mps
Node number 3	Mobility Type	Circle mobility
	Radius	100 m
	Speed	80 mps
	Start Angle	120 degrees
	mobility.cy	250 m
	mobility.cx	250 m
Node number 4	Mobility Type	Constant speed mobility
	Speed	60 mps
Node number 5	Mobility Type	Constant speed mobility
	Speed	40 mps
Node number 6	Mobility Type	Constant speed mobility
	Speed	80 mps
Node number 7	Mobility Type	Constant speed mobility
	Speed	100 mps
Node number 8	Mobility Type	Circle mobility
	Radius	120 m
	Speed	90 mps
	Start Angle	90 degrees
	mobility.cy	100 m
	mobility.cx	100 m

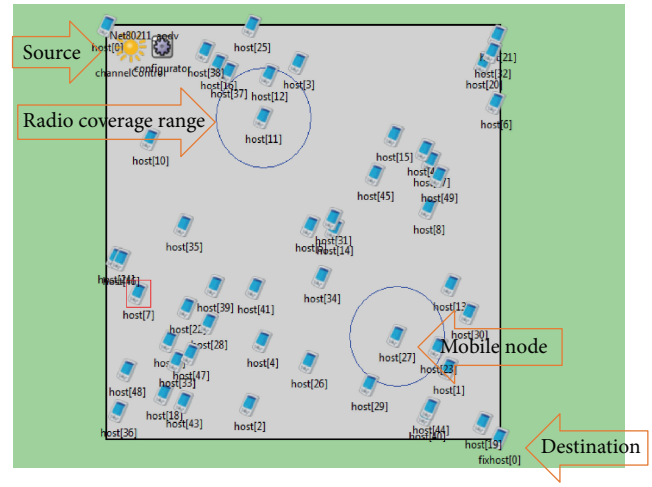


Figure 2: Initial layout of the network.

malicious nodes tend to cause more damage than stationary malicious nodes.

On the other hand, as shown in Figure 4, PASER showed very high PDR even when the network contained malicious nodes performing hello flooding attack. The highest PDR occurred when one stationary node performed the attack.

Table 7: Attacks' scenarios.

Attack	Scenario description	Scenario number
Selective forwarding and hello flooding	Network with one stationary malicious node	1
	Network with one mobile malicious node	2
	Network with two stationary malicious nodes	3
	Network with two mobile malicious nodes	4
	Network with four stationary malicious nodes	5
	Network with four mobile malicious nodes	6
	Network with eight stationary malicious nodes	7
	Network with eight mobile malicious nodes	8
Wormhole	Network where the source and the destination are stationary	9
	Network where the source is mobile and the destination is stationary	10
	Network where the source is stationary and the destination is mobile	11
	Network where the source and the destination are mobile	12

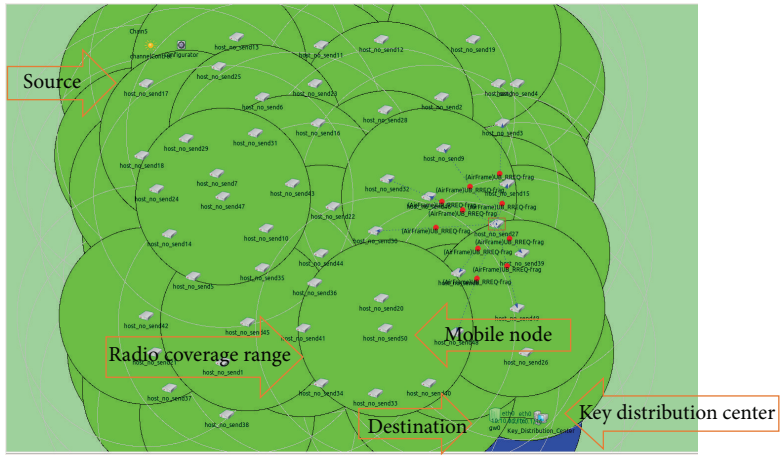


Figure 3: Radio coverage of nodes in the PASER based network.

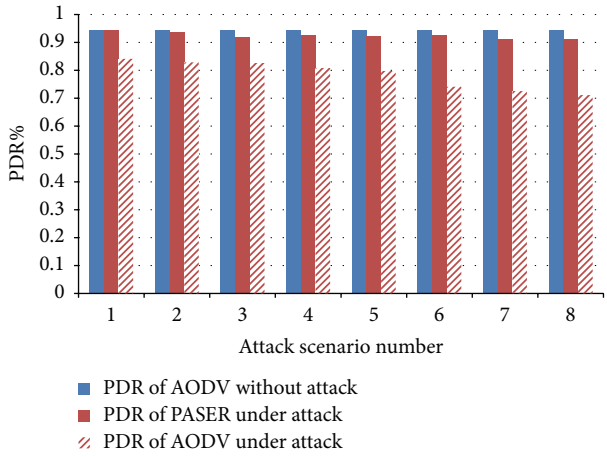


Figure 4: Impact of hello flooding attack on packet delivery ratio (PDR).

The lowest PDR was in scenario number 8; when eight malicious nodes were performing the attack, this caused only about 4.02% decrease in PDR. It can be said that PASER protocol when compared to AODV showed high PDR performance and immunity against hello flooding.

The effect of hello flooding attack on network throughput is not significant as shown in Figure 5. The largest drop was in the second case, where only one mobile node performed the attack. This decrease in throughput was only about 0.23%, which is not significant. The nature of the malicious nodes, whether they are mobile or stationary, does not affect the impact of the attack. In the first, second, fifth, and sixth attack scenarios, the mobile malicious nodes tend to cause more damage on network throughput. However, in the rest of attacks scenarios, the stationary malicious nodes caused greater decrease in network throughput. In general, network throughput is not significantly affected by hello flooding attack and mobility factor does not increase the damage caused by the attack.

The impact of implementing PASER as a security protocol appeared very clearly in network throughput, as shown in Figure 5. PASER achieved only about 56% of the throughput achieved by AODV when the two protocols were under hello flooding attack. This is because PASER uses authentication and cryptography to validate the packets, which slow down

Table 8: Results of hello flooding attacks on AODV and PASER.

Scenario number		Results with AODV	Results with PASER
1	Packet delivery ratio	84.10%	94.30%
	End-to-end delay	0.0038617 seconds	0.007586 seconds
	Throughput	9,309.219 bits/second	5,013.218 bits/second
2	Packet delivery ratio	82.86%	93.62%
	End-to-end delay	0.003969 seconds	0.007538 seconds
	Throughput	9,289.311 bits/second	5,010.701 bits/second
3	Packet delivery ratio	82.65%	92.08%
	End-to-end delay	0.004218 seconds	0.007812 seconds
	Throughput	9,308.361 bits/second	5,010.19 bits/second
4	Packet delivery ratio	80.85%	92.71%
	End-to-end delay	0.005643 seconds	0.007908 seconds
	Throughput	9,312.8620 bits/second	5,009.24 bits/second
5	Packet delivery ratio	79.80%	92.42%
	End-to-end delay	0.005712 seconds	0.007742 seconds
	Throughput	9,310.102 bits/second	5010.46 bits/second
6	Packet delivery ratio	74.13%	92.54%
	End-to-end delay	0.005811 seconds	0.007826 seconds
	Throughput	9,309.671 bits/second	5011.12 bits/second
7	Packet delivery ratio	72.54%	91.13%
	End-to-end delay	0.006105 seconds	0.007691 seconds
	Throughput	9,306.418 bits/second	5004.17 bits/second
8	Packet delivery ratio	71.15%	91.31%
	End-to-end delay	0.006412 seconds	0.007781 seconds
	Throughput	9,308.369 bits/second	4981.71 bits/second

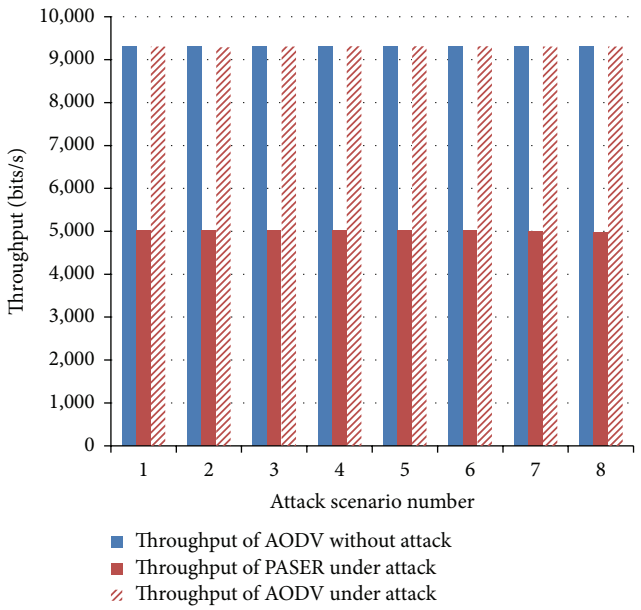


Figure 5: Impact of hello flooding attack on throughput.

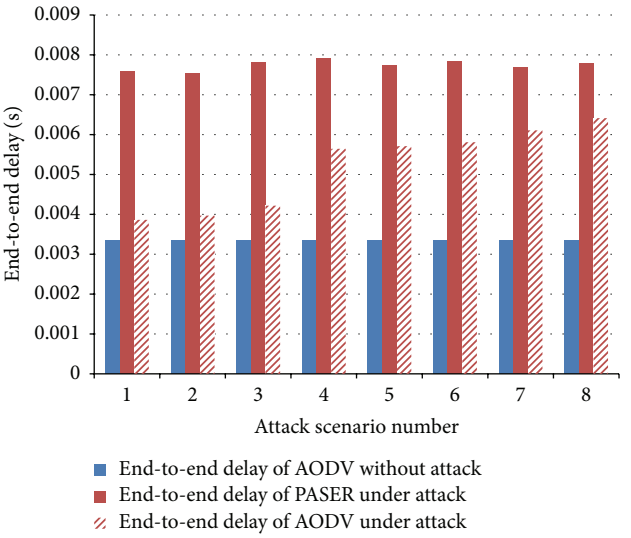


Figure 6: Impact of hello flooding attack on end-to-end delay.

the forwarding of data and routing packets. The mobility of malicious nodes did not affect the performance of PASER in protecting the network.

The impact of hello flooding attack on end-to-end delay is tangible and serious. Figure 6 shows that the effect of hello flooding attack increases when the number of malicious nodes increases, which is clear in all eight attack scenarios. Moreover, the impact of hello flooding attack on network end-to-end delay was more destructive when malicious mobile nodes performed the attack. The most significant increase in end-to-end delay took place when eight mobile malicious nodes performed the attack. In this case, the end-to-end delay increased by 91.52%, which is very high. The lowest increase in end-to-end delay took place when only one

Table 9: Results of selective forwarding attack on AODV and PASER.

Scenario number		Results with AODV	Results with PASER
1	Packet delivery ratio	87.10%	95.10%
	End-to-end delay	0.003495 seconds	0.007532 seconds
	Throughput	9,310.12 bits/second	5,012.81 bits/second
2	Packet delivery ratio	86.22%	94.64%
	End-to-end delay	0.003550 seconds	0.007698 seconds
	Throughput	9,302.11 bits/second	5,009.31 bits/second
3	Packet delivery ratio	86.13%	93.21%
	End-to-end delay	0.003701 seconds	0.007721 seconds
	Throughput	9,308.88 bits/second	5,011.45 bits/second
4	Packet delivery ratio	85.13%	91.13%
	End-to-end delay	0.00446 seconds	0.008101 seconds
	Throughput	9,301.356 bits/second	5,012.04 bits/second
5	Packet delivery ratio	83.21%	92.61%
	End-to-end delay	0.00529 seconds	0.007915 seconds
	Throughput	9,294.21 bits/second	5008.12 bits/second
6	Packet delivery ratio	82.40%	91.84%
	End-to-end delay	0.005321 seconds	0.008101 seconds
	Throughput	9,300.81 bits/second	5,010.52 bits/second
7	Packet delivery ratio	81.74%	91.72%
	End-to-end delay	0.005983 seconds	0.008106 seconds
	Throughput	9,276.47 bits/second	5,001.12 bits/second
8	Packet delivery ratio	81.13%	90.82%
	End-to-end delay	0.006018 seconds	0.008120 seconds
	Throughput	9285.18 bits/second	4,981.87 bits/second

malicious stationary node performed the attack; the increase in this situation was about 15.3%.

As Figure 6 shows, PASER caused higher end-to-end delay on the network compared to AODV when the network was under hello flooding attack. This was not an effect of the attack since the increase in the number of malicious nodes was not reflected in end-to-end delay; rather, it was associated with how PASER selects the optimal path and the time taken to validate the position of the nodes and to validate routing packets.

Since, AODV does not contain any protection mechanism and at packet loss performs route discovery. This is good in terms of throughput and end-to-end delay, but not in case of PDR. On the other hand, PASER is security protocol which is based on cryptography, mainly public/key cryptography that makes throughput and end-to-end delay of each node low, as each node has to verify the signature of the packet and then resign and forward it, but provides better PDR and thus improves the efficiency of the MWMN as shown in Figures 4, 5, and 6.

3.5. Simulation Scenarios with Malicious Nodes Performing a Selective Forwarding Attack. In this section, the effect of a selective forwarding attack on stationary and mobile networks is measured. The attack is performed by malicious nodes randomly dropping packets. The scenarios described in Section 3.3 were simulated and the collected results are shown in Table 9. The collected results then were compared to the results of AODV under the same attack.

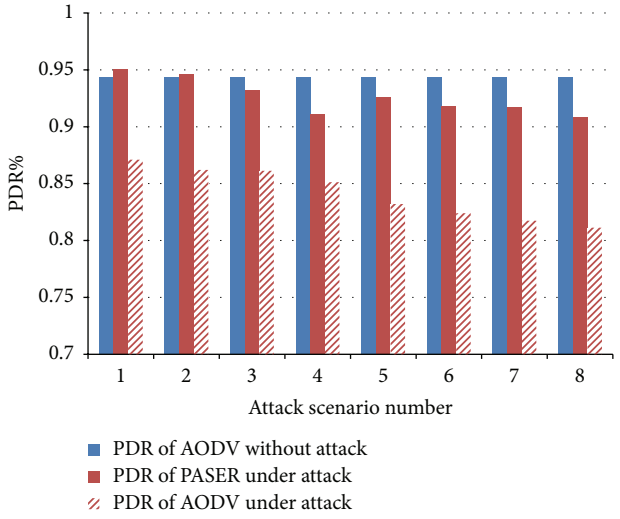


Figure 7: Impact of selective forwarding attack on packet delivery ratio (PDR).

As Figure 7 shows, even though selective forwarding attack is less harmful than hello flooding attack, it still causes significant decrease in network PDR. The worst case took place when scenario number 8 was conducted; this caused 16.36% drop in network PDR. By comparing all eight scenarios, it became clear that mobile malicious nodes introduced more damage.

Table 10: Results of wormhole attack on AODV and PASER.

Test scenario		Results with AODV	Results with PASER
9	Packet delivery ratio	84.81%	96.23%
	End-to-end delay	0.003861 seconds	0.007761 seconds
	Throughput	9,303.54 bits/second	5,012.31 bits/second
10	Packet delivery ratio	84.10%	95.71%
	End-to-end delay	0.004015 seconds	0.007532 seconds
	Throughput	9,289.51 bits/second	5,011.72 bits/second
11	Packet delivery ratio	85.90%	95.03%
	End-to-end delay	0.003928 seconds	0.007891 seconds
	Throughput	9,313.64 bits/second	5,010.68 bits/second
12	Packet delivery ratio	85.13%	94.21%
	End-to-end delay	0.003987 seconds	0.007961 seconds
	Throughput	9,301.11 bits/second	5,012.36 bits/second

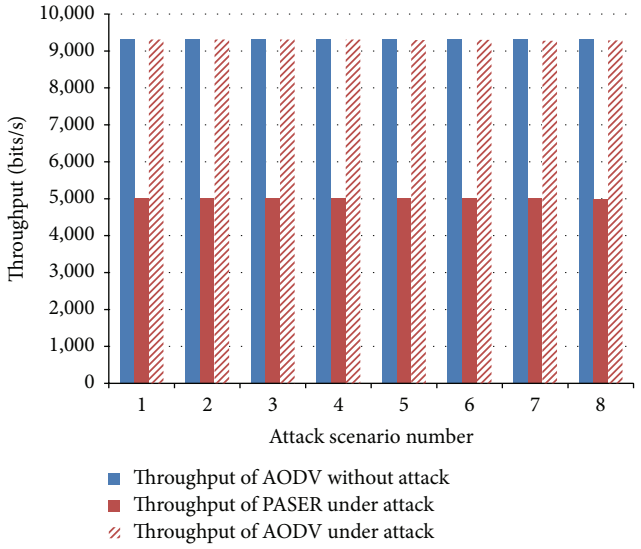


Figure 8: Impact of selective forwarding attack on throughput.

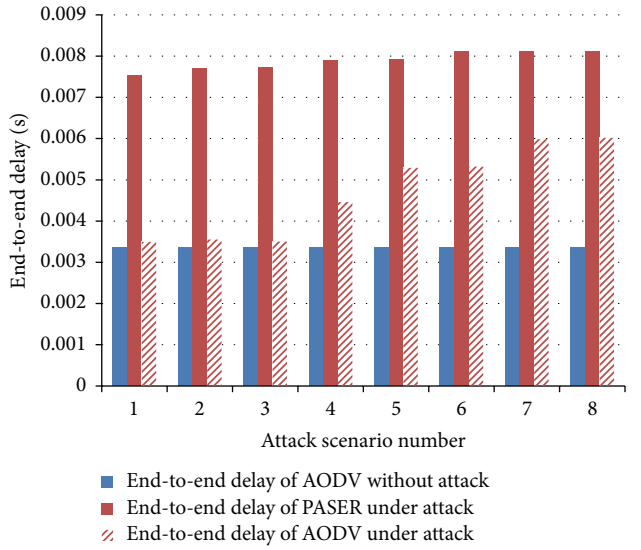


Figure 9: Impact of selective forwarding attack on end-to-end delay.

When the network was under selective forwarding attack, PASER still showed high PDR compared to AODV as shown in Figure 7. The lowest PDR for both routing protocols occurred when eight mobile malicious nodes performed the attack. In selective forwarding attack, the impact of the mobility factor was stronger, especially when eight malicious nodes performed the attack. This is clear by comparing each stationary attack scenario with mobile attack scenario for the same number of malicious nodes.

Figure 8 shows that the impact of selective forwarding attack and mobility is not significant in network throughput. The largest drop in throughput was from 9310.891 bits/second to 9276.47 bits/second when eight stationary nodes were involved in the attack. This drop was only about 0.37%.

The throughput of PASER network under selective forwarding attack was about 53.7% of AODV throughput under the same attack, as shown in Figure 8, which was approximately the same impact of hello flooding attack. Figure 8 also shows that the impact of the mobility of malicious nodes was minor.

In case of AODV as shown in Figure 9, the impact of mobile malicious nodes performing selective forwarding attack on end-to-end delay is more significant than the impact caused by stationary malicious nodes. The eighth attack scenario where the malicious nodes are mobile caused largest increase in end-to-end delay with about 79.75% increase.

End-to-end delay of PASER network under selective forwarding attack is high compared with AODV. As shown in Figure 9, end-to-end delay of PASER network was more than the double of end-to-end delay caused by AODV network under the same attack. In general, even though the difference was high, the mobile malicious nodes caused higher end-to-end delay in both AODV and PASER based networks.

3.6. Simulation Scenarios with Malicious Nodes Performing a Wormhole Attack. In this section, the impact of wormhole attack on stationary and mobile networks is measured; after conducting the attacks as described in Section 3.3, the collected results are shown in Table 10.

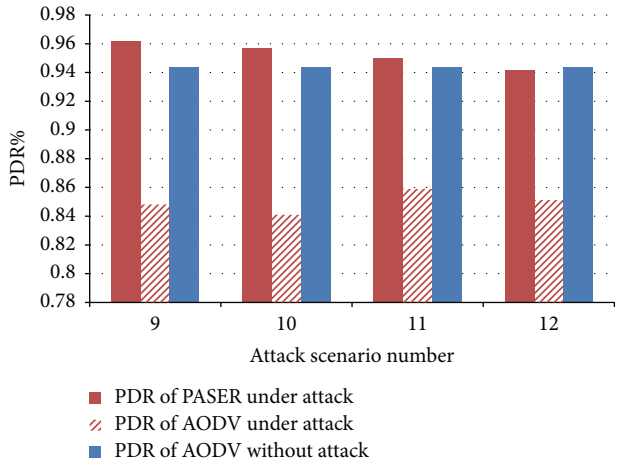


Figure 10: Impact of wormhole attack on packet delivery ratio (PDR).

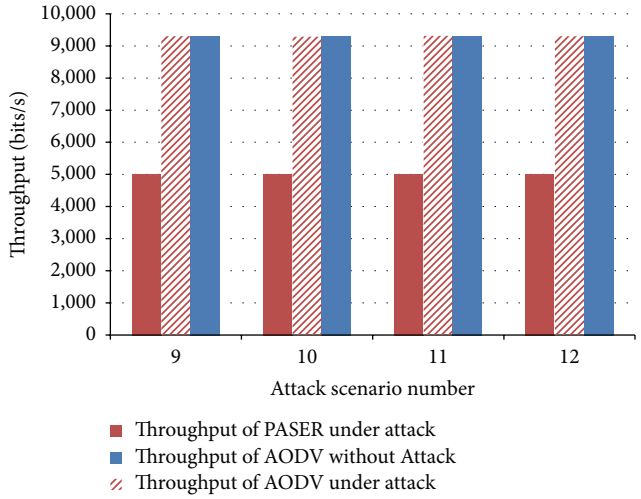


Figure 11: Impact of wormhole attack on throughput.

Figure 10 shows that the impact of wormhole attack on PDR was more significant when there is mobility in the network. The highest decrease in the PDR is about 11.22% which took place when the source of the wormhole is mobile and the destination is stationary. Moreover, Figure 10 also shows that when the source and the destination of the wormhole were stationary, the impact of the wormhole attack on the PDR was of slightly less than 10.15% drop. It can be concluded that the mobility of malicious nodes makes the impact of wormhole attack more destructive.

Moreover, as shown in Figure 10, PASER showed immunity to wormhole attack because of the usage of cryptography to authenticate packets and usage of GPS to validate nodes locations. The PDR achieved by PASER was very high compared with the PDR achieved by AODV. The highest difference in PDR between AODV and PASER took place when the source of the wormhole tunnel was stationary and the destination was mobile. In this case, PASER achieved 12.8% higher PDR. In regard to mobility, PASER was not affected by the mobility of the malicious nodes.

As Figure 11 shows, in case of AODV the impact of a wormhole attack on throughput was not significant. In the worst case, when the source of the wormhole was mobile and the destination of the wormhole was stationary, the decrease in throughput in this case was about 0.23%. Moreover, the results show that the mobility of malicious nodes did not increase the damage caused by the attack. This was found by comparing the first setup, where both ends of the tunnel were stationary with the fourth setup and where both ends of the tunnel were mobile.

In case of PASER, the throughput of the network when a wormhole attack occurs in the network was approximately the same as that with the selective forwarding and hello flooding attacks, as shown in Figure 11. AODV still achieved more than double the throughput achieved by PASER under the same type of attack. The highest difference in throughput took place in the third scenario, when the source of the wormhole was mobile and the destination was stationary. In this case, the throughput achieved by AODV was 9,313.64 bits/second, while PASER achieved 5,010.68 bits/second, a decrease of about 85.87% in network throughput.

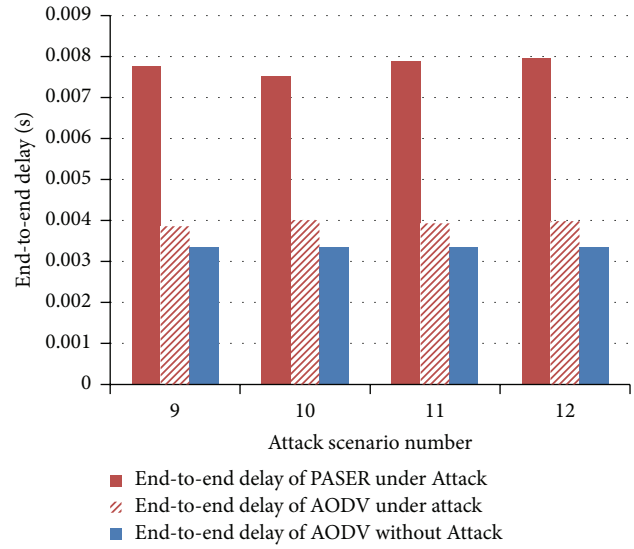


Figure 12: Impact of wormhole attack on end-to-end delay.

The results presented in Figure 12 shows that the damage caused by wormhole attack is very significant with respect to end-to-end delay. The highest increase in end-to-end delay took place when the source of the tunnel was mobile and the destination was stationary; the increase was about 19.92%. Moreover, the mobility of malicious nodes made the impact of wormhole attack more significant. This was concluded by comparing the increase in end-to-end delay when the wormhole tunnel ends were both stationary with the scenario when the wormhole tunnel ends were mobile. In the first case, the increase in end-to-end delay was about 15%, while it was about 19% in the second case.

In case of PASER, end-to-end delay in PASER network as shown in Figure 12 was affected more by wormhole attack than by selective forwarding and hello flooding attacks. The

Table 11: PDR of AODV and PASER in large networks.

Number of nodes	PDR (%) with AODV	PDR (%) with PASER
50	94.40	94.83
75	93.25	94.41
100	90.21	93.16
125	88.56	92.52
150	76.15	91.08
175	91.3	91.91
200	71.86	90.78
225	80.48	92.36
250	88.14	91.73
275	91.28	92.75
300	75.43	86.31

Table 12: Throughput of AODV and PASER in large networks.

Number of nodes	Throughput (bits/seconds) with AODV	Throughput (bits/seconds) with PASER
50	9310.891	5200.471
75	9301.413	5112.671
100	9289.21	4981.325
125	9309.77	4627.268
150	9084.18	4489.772
175	9101.46	4413.271
200	9101.83	4106.471
225	9281.93	3819.446
250	9011.71	3528.731
275	8980.76	3187.267
300	9012.14	2967.164

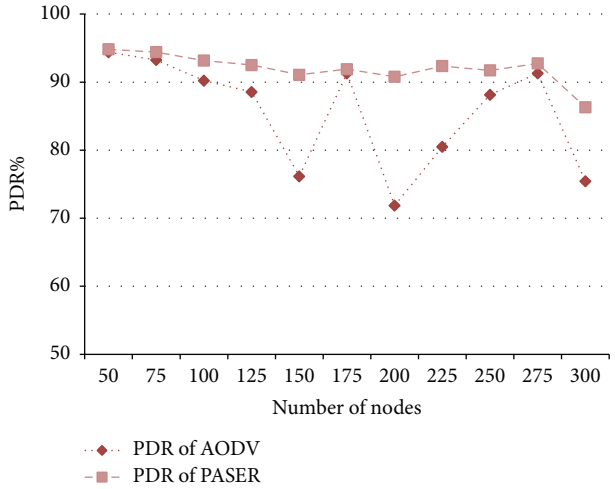


Figure 13: PDR of PASER and AODV in large networks.

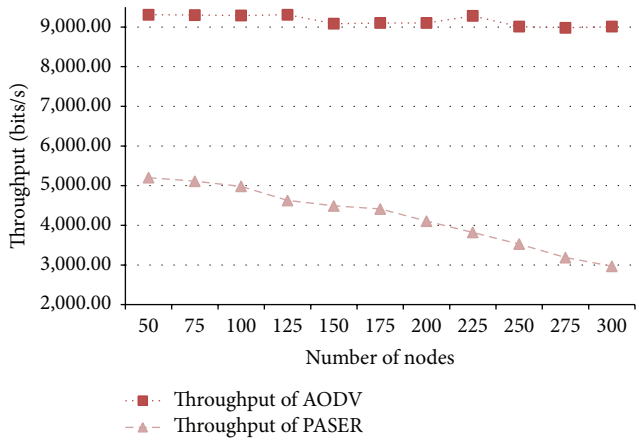


Figure 14: Throughput of PASER and AODV in large networks.

highest end-to-end delay took place when both ends of the wormhole tunnel were mobile, which is double the end-to-end delay of AODV in the same circumstances.

3.7. Performance and Scalability. After studying the impact and immunity of AODV and PASER against attacks and the impact of these attacks on performance, in this section, PASER and AODV routing protocol are tested in larger network in the same circumstances as elaborated with the number of nodes being increased from 50 to 300 nodes with increment of 25 nodes in each experiment. The attributes to be measured are PDR, end-to-end delay, and network throughput. Moreover, this section will compare the performance and scalability of both AODV and PASER.

3.7.1. Impact on Network PDR. After measuring PDR of AODV and PASER based networks while varying the nodes from 50 to 300, the collected results are listed in Table 11.

As Figure 13 shows PDR of AODV protocol varies from 94.4% to 71.86%. In the case of AODV, it can be seen that when the number of nodes increases, PDR decreases due to the increase of collision between packets when travelling from source to destination. As shown in Figure 13, PASER showed very high PDR compared to AODV in all experiments. Moreover, PASER shows very good scalability and performance of PDR when compared with AODV. The lowest reading for PASER was 86.31% when the network contains 300 nodes; the rest of reading shows that the highest network PDR drop was not more than 4.46%. On the other hand AODV PDR dropped from 94.4% to reach 71.86% when the network contains 200 nodes.

3.7.2. Impact on Network Throughput. Network throughput was measured with different network sizes for AODV and PASER as shown in Table 12. The measured values of network throughput for routing protocols are listed in Table 12.

When considering network throughput, PASER showed continuous decrease when the number of nodes increases as shown in Figure 14. The lowest throughput of PASER protocol was 2967.164 bits/second which took place when the number of nodes is 300. This is due to the increase of routing packets when the number of nodes increases, which leads to increase in signing and verification process of routing packets.

Table 13: End-to-end delay of AODV and PASER in large networks.

Number of nodes	End-to-end delay (seconds) with AODV	End-to-end delay (seconds) with PASER
50	0.003348	0.0074
75	0.003421	0.007501
100	0.003762	0.007513
125	0.003901	0.007978
150	0.003978	0.008102
175	0.003851	0.00821
200	0.004021	0.00861
225	0.004089	0.00887
250	0.004147	0.00916
275	0.004162	0.00971
300	0.004189	0.01073

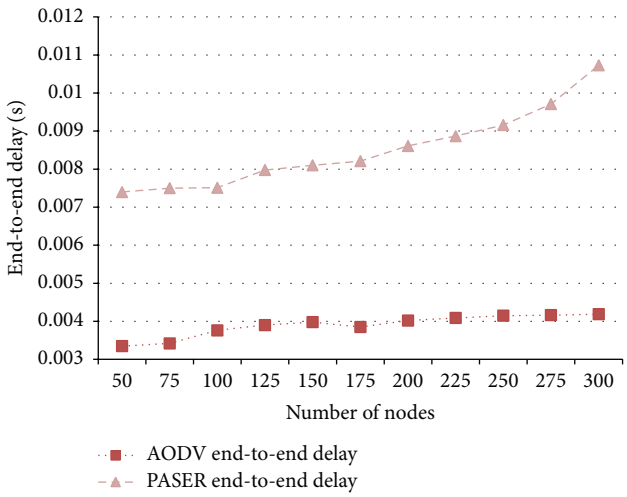


Figure 15: End-to-end delay of PASER and AODV in large networks.

3.7.3. Impact on Network End-to-End Delay. The measured values of network end-to-end delay for AODV and PASER routing protocols are listed in Table 13.

According to Figure 15, network end-to-end delay increased by 25.11% in the worst case. The lowest end-to-end delay took place when the network size was only 50 nodes with value of 0.003348 seconds, while the longest end-to-end delay was 0.004189 seconds when the number of nodes is 300. Such increase in end-to-end delay is affected by the route discovery process and number of nodes in the path from the source to destination.

The collected results of network end-to-end delay showed that PASER causes increase in end-to-end delay when the number of nodes increases. As shown in Figure 15, the increase reached about 45% in the worst case. Comparing this by AODV results shows that AODV is more scalable and can deliver better end-to-end delay performance.

4. Discussion and Recommendations

One of the goals of this research is to study the effect of hello flooding, selective forwarding, and wormhole attacks on network performance, particularly the effect on PDR, network throughput, and end-to-end delay. Table 14 shows the effect of each attack on each network parameter for both routing protocols under study. The PDR and end-to-end delay were the most affected network parameters by the attacks, while throughput was less affected. Moreover, Table 14 shows the results of using PASER to protect the network against these three types of attacks. PASER was successful in reducing the impact of all three attacks on network PDR. In regard to end-to-end delay and network throughput, PASER was able to protect the network against the attacks, but unfortunately it introduced more end-to-end delay and reduced network throughput as shown in Sections 3.4, 3.5, 3.6, and 3.7. This is due to the heavy usage of cryptography by PASER in authenticating nodes and routing packets, which led to substantial increase in end-to-end delay and decrease in network throughput.

One of the main contributions of this research is to study the impact of mobility of malicious nodes in network performance. Table 15 shows the association between the mobility of malicious nodes and level of damage caused by attack. Here, Yes means that the mobile malicious nodes caused more damage than stationary nodes, while No indicates that mobile nodes did not cause more damage than stationary nodes. As Table 15 shows, mobile nodes tended to increase the effect of attack on network PDR for AODV protocol more than in PASER except for selective forwarding attack. In regard to end-to-end delay, the effect of the attacks on both protocols increased when the malicious nodes were mobile.

The performance decrease and weak scalability of PASER protocols reveal the need for a protection mechanism to detect and isolate malicious nodes without continuous authentication. Such mechanism can be integrated into the route discovery process and used in the detection and repair of broken routes. A statistical-based methodology that uses routing history to select the safest route is an option to be considered, but it requires determining how to select the route for the first time and how the exchanged and saved routing data will affect network and nodes performance. A hybrid approach which uses cryptography and communication statistics is also worth considering. In such approach, routing protocol can use cryptography to authenticate packets and nodes to guarantee the connectivity and existence of an alternative path while gathering statistics to provide the faster path when required.

5. Conclusion and Future Work

Denial-of-service (DoS) attacks are fatal to many types of network, including wireless mesh networks, specifically when the network is utilized in a highly sensitive scenario like eHealthcare. This paper focused on studying three types of attacks that can cause DoS in static and mobile WMNs: hello flooding, selective forwarding, and wormhole attacks. The first step was to study the available solutions for DoS attacks on WMNs. The literature review indicates that there is no routing protocol that provides a comprehensive solution for the DoS problem in WMNs. In general, the proposed solutions tend to deal with a specific attack or a group of

Table 14: Impact of all attacks on network parameters.

Routing protocol	Attack	Network parameter		
		PDR	Throughput	End-to-end delay
AODV	Hello flooding	−32.68%	−0.23%	+91.52%
	Selective forwarding	−16.36%	−0.37%	+79.75%
	Wormhole	−11.22%	−0.23%	+19.92%
PASER	Hello flooding	−4.02%	−0.94%	+6.86%
	Selective forwarding	−4.10%	−0.93%	+9.73%
	Wormhole	−0.67%	−0.35%	+7.58%

Table 15: Impact of mobility of malicious nodes on AODV and PASER.

Routing protocol	Attack	Network parameter		
		PDR	Throughput	End-to-end delay
AODV	Hello flooding	Yes	No	Yes
	Selective forwarding	Yes	No	Yes
	Wormhole	Yes	No	Yes
PASER	Hello flooding	No	No	Yes
	Selective forwarding	Yes	No	Yes
	Wormhole	No	No	Yes

attacks without being able to secure the routing layer completely. Most of the proposed solutions utilize cryptography to authenticate routing packets and neighboring nodes, which assumes that nodes can register with an authority to obtain certificates and validate themselves. Moreover, it assumes that these nodes will not perform malicious activities. The second proposed methodology is to use past communication statistics to identify malicious nodes that are trying to disrupt the network. Hardware-based solutions are also suggested to solve wormhole attacks.

To implement the attacks and measure their effect, it was necessary first to select the routing protocol to be used in the network layer of a WMN. AODV, a very common routing protocol, was selected to be analyzed, and its immunity to the effect of these attacks on the packet delivery ratio, throughput, and end-to-end delay of the network was measured. To perform this task, OMNET++, a simulation tool, was selected to implement the three types of attacks and launch them against a mobile WMN. In general, the results showed that the significantly affected network performance parameters were the PDR and throughput.

After determining the impact of the attacks on AODV, the next step was to decide which security protocol to use as defense protocol against the attacks. One of the mature proposed solutions selected was Position-Aware, Secure, and Efficient Mesh Routing (PASER). PASER routing protocol depends heavily on public key cryptography to secure routing by authenticating routing packets and nodes. Moreover, PASER uses a GPS module to validate nodes' locations to detect wormhole tunnels and prevent attacks. The results show that PASER is very effective in protecting the network against attacks, but, unfortunately, it imposes a performance cost on the PDR and end-to-end delay because of the heavy usage of cryptography. Moreover, PASER assumes the existence of a key distribution center (KDC) to issue keys for the nodes, which might not always be possible, and prevents the dynamic enrollment of nodes in the WMN.

During the process of writing this paper, it became clear that there is room for future improvement and enhancement, as the literature lacks solid implementation and simulation modules for all types of attacks to be used in testing and evaluating new proposed solutions. Providing such a framework will help researchers to provide practical solutions. Reviewing the proposed solutions reveals that designing a smart detection mechanism that does not depend on cryptography to detect and isolate malicious nodes is very helpful and will help in advancing WMN security research. In the future, these types of mechanisms and protocols for detecting malicious nodes can be investigated, and their effectiveness in protecting the network can be studied.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This project was funded by the National Plan of Science, Technology and Innovation (MAARIFAH), King Abdulaziz City for Science and Technology, Saudi Arabia (Award no. 12-INF2817-02).

References

[1] I. F. Akyildiz, X. Wang, and W. Wang, "Wireless mesh networks: a survey," *Computer Networks*, vol. 47, no. 4, pp. 445–487, 2005.

[2] P. H. Pathak and R. Dutta, "A survey of network design problems and joint design approaches in wireless mesh networks," *IEEE Communications Surveys & Tutorials*, vol. 13, no. 3, pp. 396–428, 2011.

[3] S. D. Odabasi and A. H. Zaim, "A survey on wireless mesh networks, routing metrics and protocols," *International Journal of Electronics, Mechanical and Mechatronics Engineering*, vol. 2, no. 1, pp. 92–104, 2010.

[4] C.-W. Lee, "Security in wireless mesh networks," in *Wireless Network Security*, pp. 229–246, Springer, Berlin, Germany, 2013.

[5] R. Matam and S. Tripathy, "WRSR: wormhole-resistant secure routing for wireless mesh networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2013, no. 1, article 180, 2013.

[6] Z. You and Y. Wang, "An efficient and secure routing protocol for a hybrid wireless mesh network," *Journal of Computational Information Systems*, vol. 8, no. 21, pp. 8693–8705, 2012.

[7] S. Khan and J. Loo, "Cross layer secure and resource-aware on-demand routing protocol for hybrid wireless mesh networks," *Wireless Personal Communications*, vol. 62, no. 1, pp. 201–214, 2012.

[8] S. Khan, N. A. Alrajeh, and K.-K. Loo, "Secure route selection in wireless mesh networks," *Computer Networks*, vol. 56, no. 2, pp. 491–503, 2012.

[9] D. Benetti, M. Merro, and L. Viganò, "Model checking ad hoc network routing protocols: ARAN vs. endair A," in *Proceedings of the 8th IEEE International Conference on Software Engineering and Formal Methods (SEFM '10)*, pp. 191–202, September 2010.

[10] S. Lu, L. Li, K.-Y. Lam, and L. Jia, "SAODV: a MANET routing protocol that can withstand black hole attack," in *Proceedings of the International Conference on Computational Intelligence and Security (CIS '09)*, pp. 421–425, Beijing, China, December 2009.

[11] C. H. Lin, W. S. Lai, Y. L. Huang, and M. Chou, "I-SEAD: a secure routing protocol for mobile Ad Hoc networks," *Multimedia and Ubiquitous Engineering*, vol. 1, no. 1, pp. 102–107, 2008.

[12] M. O. Pervaiz, M. Cardei, and J. Wu, "Routing security in Ad Hoc wireless networks," *Network Security*, pp. 117–142, 2010.

[13] K. Graffi, P. S. Mogre, M. Hollick, and R. Steinmetz, "Detection of colluding misbehaving nodes in mobile ad hoc and wireless mesh networks," in *Proceedings of the 50th Annual IEEE Global Telecommunications Conference (GLOBECOM '07)*, pp. 5097–5101, Washington, DC, USA, November 2007.

[14] S. Madria and J. Yin, "SeRWA: a secure routing protocol against wormhole attacks in sensor networks," *Ad Hoc Networks*, vol. 7, no. 6, pp. 1051–1063, 2009.

[15] H.-M. Sun, C.-H. Chen, C.-W. Yeh, and Y.-H. Chen, "A collaborative routing protocol against routing disruptions in MANETs," *Personal and Ubiquitous Computing*, vol. 17, no. 5, pp. 865–874, 2013.

[16] M. Sbeiti, A. Wolff, and C. Wietfeld, "PASER: Position aware secure and efficient route discovery protocol forwireless mesh networks," in *Proceedings of the 5th International Conference on Emerging Security Information, Systems and Technologies (SECURWARE '11)*, pp. 63–70, Saint Laurent du Var, France, August 2011.

[17] A. Sgora, D. D. Vergados, and P. Chatzimisios, "A survey on security and privacy issues in wireless mesh networks," *Security and Communication Networks*, 2013.

[18] J. Sen, "Security and privacy issues in wireless mesh networks: a survey," in *Wireless Networks and Security*, Signals and Communication Technology, pp. 189–272, Springer, Berlin, Germany, 2013.

[19] S. Alanazi, J. Al-Muhtadi, A. Derhab et al., "On resilience of Wireless Mesh routing protocol against DoS attacks in IoT-based ambient assisted living applications," in *Proceedings of the 17th International Conference on E-Health Networking, Application & Services (HealthCom '15)*, pp. 205–210, IEEE, Boston, Mass, USA, October 2015.

[20] S. M. S. Bari, F. Anwar, and M. H. Masud, "Performance study of hybrid Wireless Mesh Protocol (HWMP) for IEEE 802.11s WLAN mesh networks," in *Proceedings of the International Conference on Computer and Communication Engineering (ICCCE '12)*, pp. 712–716, July 2012.

[21] M. Shivlal and S. Kumar, "Performance analysis of secure wireless mesh networks," *Research Journal of Recent Sciences*, vol. 1, no. 3, pp. 80–85, 2012.

[22] T.-M. Hoang, V.-L. Dinh, and K.-Q. Nguyen, "A study on routing performance of 802.11 based wireless mesh networks under serious attacks," in *Proceedings of the IEEE International Conference on Computing, Management and Telecommunications (ComManTel '13)*, pp. 295–297, Ho Chi Minh City, Vietnam, January 2013.

[23] S. Bhumireddy, S. Tripathy, and R. Matam, "Secure peer-link establishment in wireless mesh networks," *Advances in Intelligent Systems and Computing*, vol. 176, no. 1, pp. 189–198, 2012.

[24] C. Perkins, E. Belding-Royer, and D. Samir, *Ad Hoc on Demand Distance Vector (AODV) Routing (RFC 3561)*, IETF MANET Working Group, 2003.

[25] P. Owczarek and P. Zwierzykowski, "Review of simulators for wireless mesh networks," *Journal of Telecommunications & Information Technology*, vol. 2014, no. 3, pp. 82–89, 2014.

[26] B. Kannhavong, H. Nakayama, Y. Nemoto, N. Kato, and A. Jamalipour, "A survey of routing attacks in mobile ad hoc networks," *IEEE Wireless Communications*, vol. 14, no. 5, pp. 85–91, 2007.

[27] N. S. Chouhan and S. Yadav, "Flooding attacks prevention inMANET," *International Journal of Computer Technology and ElectronicsEngineering (IJCTEE)*, vol. 1, no. 3, 2011.

[28] M. O. Khozium, "Hello flood countermeasure for wireless sensor networks," *International Journal of Computer Science and Information Security*, vol. 2, no. 3, pp. 57–65, 2008.

[29] U. Khartad and R. K. Krishna, "Route request flooding attack using trust based security scheme in Manet," *International Journal of Smart Sensors and Ad Hoc Networks*, vol. 1, no. 4, pp. 27–33, 2012.

[30] D. M. Shila, Y. Cheng, and T. Anjali, "Mitigating selective forwarding attacks with a channel-aware approach in WMNS," *IEEE Transactions on Wireless Communications*, vol. 9, no. 5, pp. 1661–1675, 2010.

[31] P. Zhou, Z. Xiang, and Y. Chen, "Detection method of gray-hole node in wireless mesh networks," in *Proceedings of the 5th International Conference on Computational and Information Sciences (ICCIS '13)*, pp. 1570–1573, IEEE, Shiyang, China, June 2013.

[32] L. Lazos and M. Krunz, "Selective jamming/dropping insider attacks in wireless mesh networks," *IEEE Network*, vol. 25, no. 1, pp. 30–34, 2011.

[33] S. Khanam, H. Y. Saleem, and A. K. Pathan, "An efficient detection model of selective forwarding attacks in wireless mesh networks," in *Internet and Distributed Computing Systems*, pp. 1–14, Springer, Berlin, Germany, 2012.

[34] A. A. Bhosle, T. P. Thosar, and S. Mehatre, "Black-hole and wormhole attack in routing protocol AODV in MANET," *International Journal of Computer Science, Engineering and Applications (IJCSEA)*, vol. 2, no. 1, pp. 45–54, 2012.

[35] R. Maulik and N. Chaki, "A study on wormhole attacks in MANET," *International Journal of Computer Information Systems and Industrial Management Applications*, vol. 3, no. 2, pp. 271–279, 2011.

[36] J. D. Parmar, A. D. Patel, R. H. Jhaveri, and B. I. Shah, "MANET routing protocols and wormhole attack against AODV," *International Journal of Computer Science and Network Security*, vol. 10, no. 4, pp. 12–18, 2010.

Hindawi Publishing Corporation
Mobile Information Systems
Volume 2016, Article ID 6850168, 10 pages
http://dx.doi.org/10.1155/2016/6850168

Research Article

Integrated Wearable System for Monitoring Heart Rate and Step during Physical Activity

Eka Adi Prasetyo Joko Prawiro,[1] Chun-I Yeh,[1] Nai-Kuan Chou,[2] Ming-Wei Lee,[1] and Yuan-Hsiang Lin[1]

[1]*Department of Electronic and Computer Engineering, National Taiwan University of Science and Technology, Taipei 10607, Taiwan*
[2]*Department of Surgery, National Taiwan University Hospital, Taipei 10048, Taiwan*

Correspondence should be addressed to Yuan-Hsiang Lin; linyh@mail.ntust.edu.tw

Received 29 January 2016; Accepted 17 April 2016

Academic Editor: Basit Shahzad

This paper integrates a heart rate (HR) monitoring system with step counter for use during physical activities. Novel step counter algorithm has been developed to enable the highly accurate detection of step. The proposed system comprises a wireless wearable device, a smartphone, and a remote server. Data transmission between a wearable device and a smartphone is conducted via Bluetooth low energy (BLE). An indirect contact measurement method has also been devised to eliminate the need for direct contact electrodes and likelihood of skin irritation. The proposed system is compact, lightweight, and comfortable to wear. A smartphone application provides the interface for the display of data related to HR, step count (SC), exercise intensity, speed, distance, and calories burned, as well as waveforms related to ECG and step cycle. ECG peak detection algorithm achieved accuracy of 99.7% using the MIT-BIH ST Change Database. Accuracy of 98.89% was achieved for HR and 98.96% for SC at treadmill speeds of 1.8 to 9.0 km/h.

1. Introduction

Global annual expenditure in the health-related domain is currently US$5.3 trillion, and this is expected to increase at an ever-accelerating rate [1]. Evidence supports a strong link between regular physical exercise and decreased risk of cardiovascular disease [2] as well as chronic illnesses such as diabetes [3].

It has been suggested that the volume of exercise required by most individuals is equivalent to 10,000 steps per day [4]. Pedometers are sensors that are worn on the body to motivate one in the pursuit of physical activity and assess one's progress. Pedometers are viewed as a practical alternative to group health promotion because the output (e.g., steps taken, steps/day) is user-friendly [5]. Unfortunately, most pedometers give no indication of the intensity of the physical activity, despite the importance of meeting minimum exertion thresholds and the danger of exceeding physical limits. For example, 55–90% of maximum HR has been cited as an appropriate level for training cardiorespiratory fitness without leading to early fatigue [6]. This underlines the need to monitor exercise intensity in real time. HR can be used as an indicator of exercise intensity [7], according to the formula outlined by Fox and Haskell [8].

Sun and Yu [9] designed an HR monitoring and fatigue detection system for drivers. Khan et al. [10] enabled the monitoring of HR using photoplethysmographics (PPG). He et al. [11] designed an ear-worn HR monitoring system using a combination of ECG, PPG, and ballistocardiogram (BCG). Accelerometers have also been used to monitor activity. Gupta and Dallas [12] developed an activity monitoring system using a single triaxial accelerometer. Cheng et al. [13] proposed a system for activity monitoring and fall detection using accelerometer signals and Ryu et al. [14] proposed a step counter using accelerometer data. Nonetheless, a comprehensive system for the evaluation of one's overall physical condition in real time has yet to be developed.

A shift toward technology that prioritizes comfort for the user has led to the development of systems based on coupling capacitance for indirect contact ECG in order to

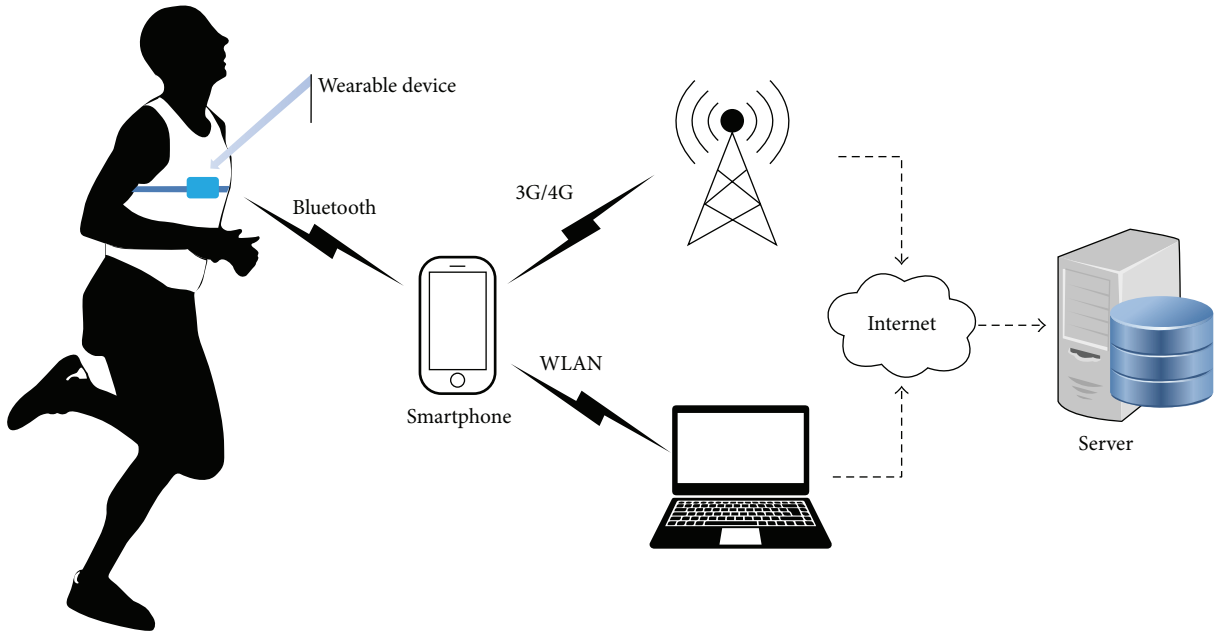


Figure 1: System architecture.

monitor HR. Oehler et al. [15] first used capacitive sensors to measure multichannel ECGs. Matsuda and Makikawa [16] implemented ECG capacitive sensors in driving applications. In 2012, Baek et al. [17] embedded in a smart chair indirect contact ECG for the measurement of physiological signals. Eilebrecht et al. [18] proposed a measurement system using capacitive multichannel ECGs. However, no previous study has conducted a comprehensive investigation into the monitoring of everyday activities such as walking and running.

Smartphones have recently been adapted to accommodate p-health monitoring systems [19, 20] and rapid progress in this area makes it a particularly useful medium for mobile monitoring systems. In this study, we developed a wearable device integrating an accelerometer-based pedometer with single-channel ECG system for the monitoring of HR, SC, exercise intensity, speed, distance, and calories burned as well as the presentation of ECG waveforms and depictions of step cycle during physical activities. We also developed a novel step counter algorithm to enable the highly accurate detection of step. The proposed device is strapped to the chest using indirect contact to reduce the likelihood of skin irritation.

2. Methods

2.1. System Architecture. The proposed monitoring system consists of three parts: a wireless wearable device, a smartphone, and a remote server. Figure 1 illustrates the architecture of the system. The wearable device includes two-electrode circuits for the measurement of ECG signals, a 3-axis accelerometer for the measurement of step signals, and an ultra-low-power microcontroller (MSP430) for data acquisition and calculations. Data transmission between the wearable device and the smartphone is conducted via BLE [21]. A smartphone application was developed to receive user data and display the HR, step count, exercise intensity, speed, distance, and calorie burn information. The sensors data can be transmitted to the smartphone in real time; the ECG and step signal waveforms can also be displayed on the screen of smartphone. If the system detects an instance of overexercise, the smartphone triggers an alarm (sound and vibration) to warn the user. At the same time, HR, SC, and GPS location are transmitted to a remote server (remote health care center) via a 3G mobile network, to enable clinical staff to monitor the activity status of the user.

2.2. Wearable Device. The wearable device comprises three parts, sensors and circuits, microcontroller, and wireless data transmission interface, as shown in the block diagram in Figure 2. The sensors and circuits include an accelerometer and ECG circuits that use I2C and ADC on a microcontroller. Wireless data transmission is achieved using a CC2541 module with BLE support. All external circuits are implemented using SMD to reduce the size of the hardware. Figure 3 presents a prototype of the wearable device, which is compact (6×4 cm) and lightweight (19 g including a 110 mAh battery).

2.2.1. Sensors, Circuits, and Microcontroller. The sensors include POLAR belt electrodes for single-channel ECG acquisition and a 3-axis accelerometer (ADXL 345) for the detection of body movement.

The main chip in the wearable device is the MSP430, which includes an ultra-low-power microprocessor integrating memory, MCU, ADC, I2C, UART data communication, and other peripheral functions on a single chip [22]. For the ECG circuits, we modified a simple, low-cost circuit using the two-electrode nondifferential amplifier proposed by Dobrev et al. [23], a block diagram of which is presented in Figure 4. The ECG signals first pass through a biopotential amplifier

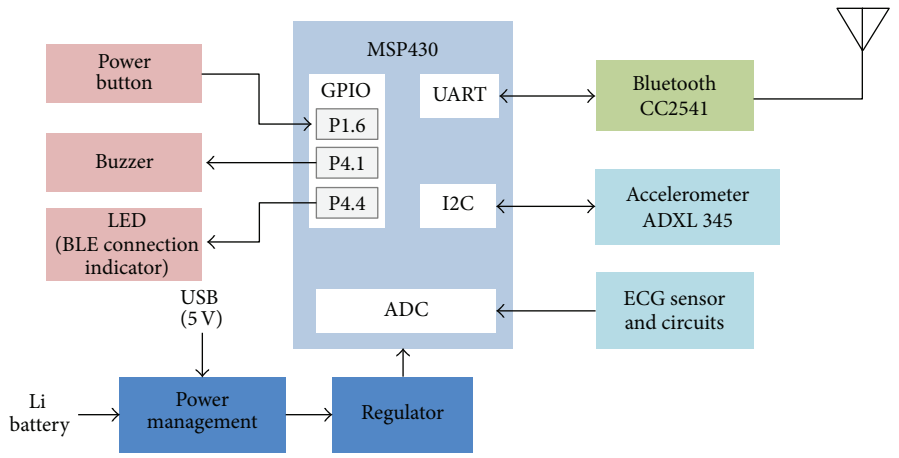
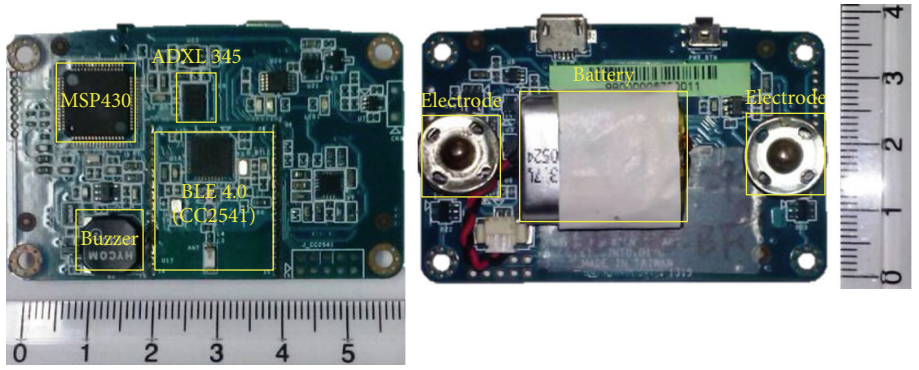


Figure 2: Block diagram of wireless wearable device.



(a)

(b)

Figure 3: Prototype wearable device: (a) left-top and right-top point of view and (b) device with cover and belt.

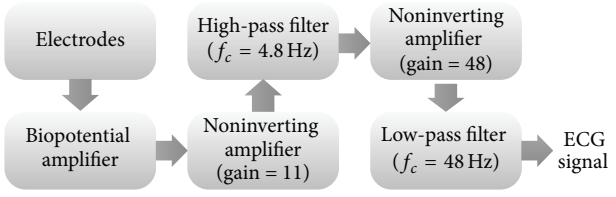


Figure 4: Block diagram of ECG circuits.

and noninverting amplifier with gain of 11, followed by a high-pass filter and noninverting amplification with cutoff frequency of 4.8 Hz and gain of 48. Finally, a low-pass filter with cutoff frequency at 48 Hz is applied.

We also employed a digital accelerometer (ADXL 345) with 13-bit resolution and acceleration range of ±16 g. The sampling rates of the ECG and accelerometer are 200 Hz and 50 Hz, respectively. The MSP430 processes the digital data to calculate the HR and SC, whereupon the processed data is sent to UART connected to a BLE module for data transmission at a data rate of 115.2 kb/s.

2.2.2. Wireless Data Transmission. The wearable device uses a standard BLE 4.0 as a data transmission interface, due to its

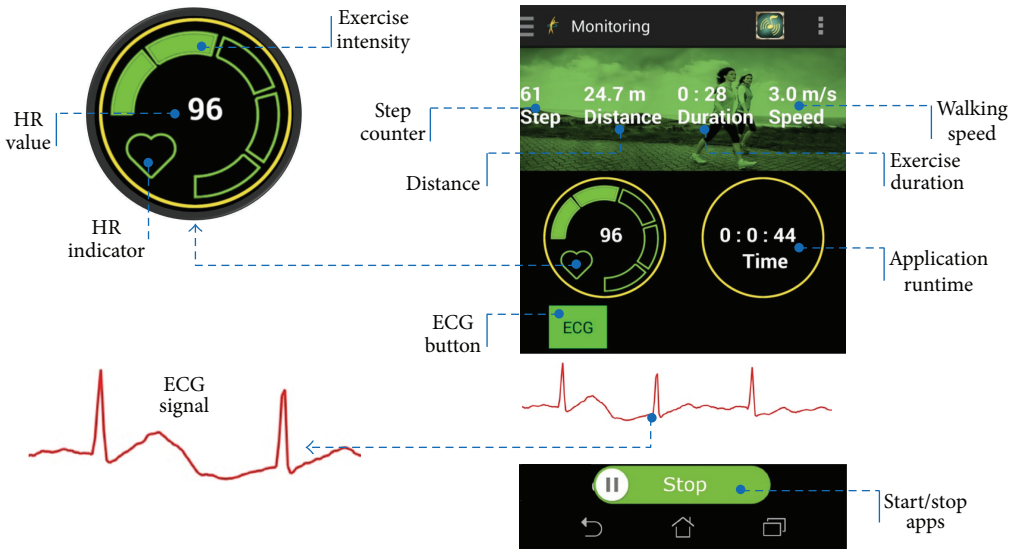


Figure 5: GUI on smartphone showing the various functions.

wide compatibility with mobile devices such as smartphones. We use CC2541 Bluetooth module [24], which is Bluetooth 4.0 that has low power energy and appropriate for healthcare applications.

2.3. Smartphone Platform. We used the ASUS Padfone S with an Android 4.4.2 operating system in conjunction with Android Eclipse for the development of a smartphone application, which we entitled "*Integrated Monitoring System.*" This application makes it possible to save HR, SC, and related parameters on a memory card and display it on the screen of the smartphone. It also initiates call to emergency response services should the need arise. Figure 5 presents the graphic user interface (GUI) on the smartphone and the function of each component.

2.4. HR Detection Algorithm. We developed a heart rate detection algorithm based on our previous work in [25]. This algorithm is divided into two stages: preprocessing and decision. Preprocessing includes a band-pass filter, differentiation, absolute function, and moving window integration (MWI). The output from this stage is then used as the input for the decision stage in which HR is calculated. The HR value is used to derive exercise intensity (1) with a corresponding maximum HR for each subject (2). For example, using 24 years as a default age value, we obtain the maximum HR as follows:

$$\text{Exercise\_Intensity} = \frac{\text{HR}}{\text{Max\_HR}} \times 100\%, \tag{1}$$

$$\text{Max\_HR} = 220 - \text{Age}. \tag{2}$$

2.5. Step Counter Algorithm. The step counter algorithm is based on that of a pedometer, wherein the number of steps can be determined after one's cycle is recognized. The accelerometer is highly sensitive to small motions; therefore, we applied several preprocessing steps to adjust the x-axis data ($X(n)$) of the signal from the accelerometer, where n is the number of data samples. Preprocessing includes band-pass filtering, the removal of baseline wandering, obtaining a moving average, enlargement, and integration. Figure 6 presents the workflow involved in preprocessing the signal, beginning with the signal passing through an FIR digital band-pass filter with cutoff frequencies of 0.5 Hz and 2 Hz. These cutoff frequencies are in accordance with the method proposed by Libby [26], with modifications to reduce noise. Baseline wandering removal (X_{subdc}) is used to remove the DC offset from the band-pass filter output (X_{band}). This is achieved by reducing each data sample using the average data point. The number of Len is 43, and j represents data number from $n-\text{Len}-1$ to $n-1$. Third, we implement an 8-tap moving average filter to smooth the signal (X_{ma}). The formulations of baseline wandering removal and moving average are given in (3) and (4), respectively. Consider

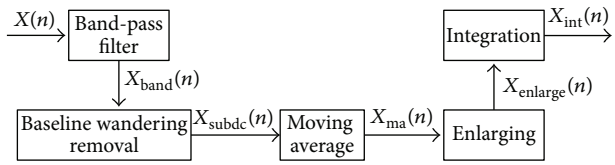


Figure 6: Workflow involved in signal preprocessing.

$$X_{\text{subdc}}(n) = X_{\text{band}}(n) - \frac{\sum_{j=n-\text{Len}-1}^{n-1} X_{\text{band}}(j)}{\text{Len}}, \tag{3}$$

$$X_{\text{ma}}(n) = \frac{\sum_{j=n-N}^{n} X_{\text{subdc}}(j)}{N}. \tag{4}$$

The moving average filter is a simple low-pass filter. The cutoff frequency in terms of moving average is given by (5), where f_s, f_c, and N are the sampling frequency (50 Hz), cutoff frequency, and number of taps (8 taps), respectively. Consider

$$f_c = 0.443\left(\frac{f_s}{N}\right). \tag{5}$$

Equation (5) is obtained from (6), where N is the number of filter taps and f is equal to f_c/f_s. Consider

$$H[f] = \frac{\sin(\pi f N)}{N\sin(\pi f)}, \quad \text{when } H[f] \text{ is } -3\,\text{dB}. \tag{6}$$

An enlarging process (X_{enlarge}) is also used to enlarge the signal characteristics using (7), as follows:

$$\begin{aligned} X_{\text{enlarge}}(n) &= \left(\frac{X_{\text{ma}}(n)}{10}\right)^2, \quad X_{\text{ma}}(n) \geq 0, \\ X_{\text{enlarge}}(n) &= X_{\text{ma}}(n), \quad X_{\text{ma}}(n) < 0. \end{aligned} \tag{7}$$

Finally, an 8-tap (N) integration process (X_{int}) is then applied to enlarge and smooth, where j represents the number of data points from $n - N$ to n, as follows:

$$X_{\text{int}}(n) = \sum_{j=n-N}^{n} X_{\text{enlarge}}(j). \tag{8}$$

Figure 7 illustrates the algorithm used in step detection. As shown in Figure 8(b), a fall is defined as a situation in which the amplitude of the current input is lower than the previous input (input(n) < input($n - 1$)). A maximum value is set when the first fall signal is detected, and a minimum value is set when current input is lower than previous one. The maximum and minimum values are used to detect peaks and peak amplitudes, which are then used to determine the interval threshold. After a peak is detected, the peak-to-peak interval is compared with the interval threshold, adapted from an experience-based rule (Table 1). Situations in which the peak-to-peak interval is higher than the interval threshold are recognized as step, whereupon the maximum and minimum values are reset and the process is repeated from the beginning. Remember that the amplitude of the step signal is associated with the speed walking; that is, a slower walking speed has a smaller amplitude and longer interval and vice versa. This adaptive threshold is used to increase the accuracy of step detection for various speeds.

2.6. Experiment Setup. To verify the performance of ECG peak detection, we use the open-source MIT-BIH ST Change Database [27] as input data. We also compared the reliability of the ECG acquisition device using *CardioSoft* [28]. To evaluate the performance of the HR and SC algorithm, we recruited five healthy males to walk and run on a treadmill. The participants ranged in age from 22 to 28 years (mean: 24 years, SD: 2.4 years), with body mass ranging from 63 to 79 kg (mean: 69 kg, SD: 6.7 kg). Measurements were obtained sequentially at the following treadmill speeds: 1.8, 2.7, 3.6, 4.5, 5.4, 6.3, 7.2, 8.1, and 9 km/h. Measurements were taken for a period of one minute at each speed, before increasing the speed by 0.9 km/h and repeating the process. Subjects wore thin clothes with the device fastened across the chest with a belt. The subjects walked for the first six speeds and ran for the last three speeds. HR data was recorded for two seconds at each speed using the proposed device as well as a commercial device (*GE Dash 3000*) [28] for comparison. The SC results from the proposed device were compared with those obtained from two commercial products (*Omron HJ-720* and *Yamax SW200*), the accuracy and reliability of which are proven [29, 30]. *HJ-720* detects step using dual accelerometers, whereas the *SW200* detects step using a traditional mechanical method. The subjects and an observer were also requested to count the number of steps taken at each speed for use as a standard (actual number of steps).

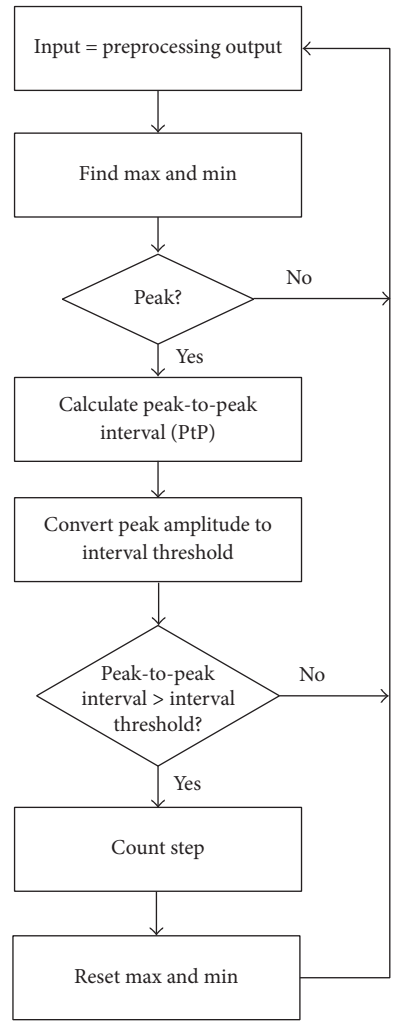


Figure 7: Step detection algorithm.

2.7. Evaluation Methods. To verify the reliability of the ECG acquisition device, we began by comparing the ECG signal from the proposed device with that of the commercial device, *CardioSoft*. We then applied our ECG R-peak detection

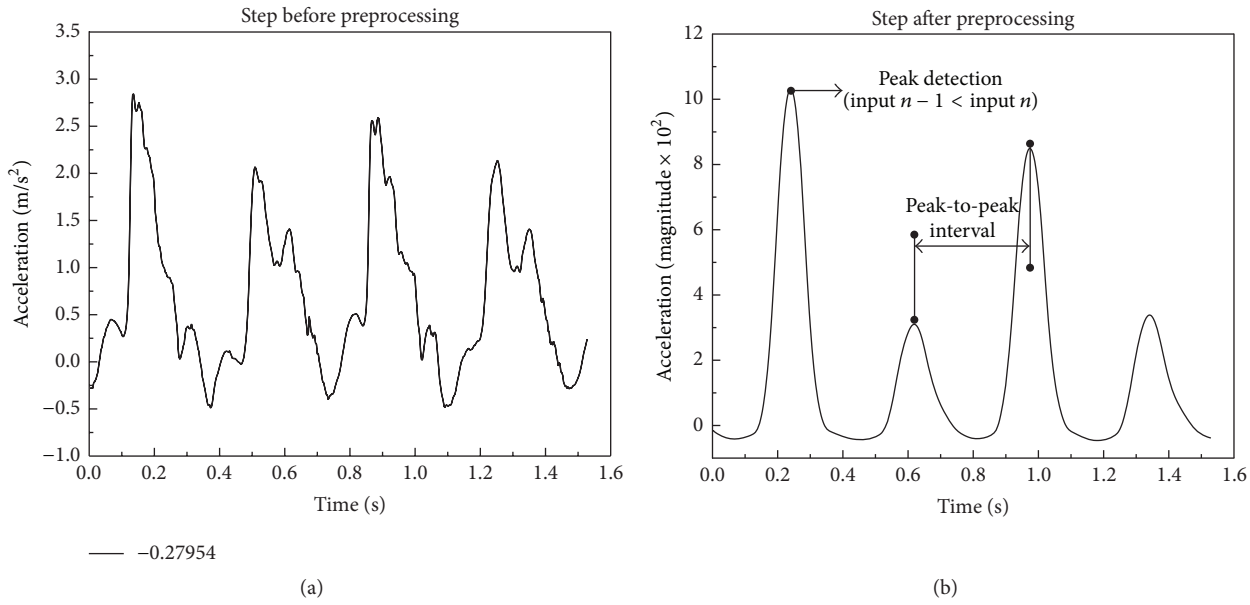


Figure 8: Step signal: (a) before preprocessing and (b) peak detection.

Table 1: Relationship between peak-to-valley amplitude and adaptive threshold.

Peak amplitude (m/s^2)	Peak-to-valley amplitude	Adapt. THR. (number of samples (second))	Walking status
<1.25	<50		Resting
1.25~1.36	50~200	22 (0.44)	Walking
1.36~1.44	200~300	20 (0.40)	
1.44~1.50	300~500	17 (0.34)	
1.50~2.03	500~650	15 (0.30)	↓
2.03~2.25	650~850	13 (0.26)	
2.25~2.61	850~1100	11 (0.22)	
2.62~2.82	1100~1600	10 (0.20)	
>2.82	>1600	9 (0.18)	Running

algorithm to 26 records from the open-source MIT-BIH ST Change Database. The aim was to identify the number of beats identified by the proposed algorithm as False Positives (FPs) and False Negatives (FNs) and then calculate the accuracy based on Real Beats (RBs) obtained from the database. The equation used to determine the accuracy of ECG peak detection is given in (9). We then applied HR detection to the data acquired from the five subjects in the treadmill test. HR accuracy was based on the HR from *GE Dash 3000* as a reference. We compared the accuracy of HR detection using the proposed device and a GE Dash 3000 in two-second increments over a period of one minute at each of the speeds between 1.8 and 9.0 km/h:

$$\text{Accuracy} = \left(1 - \left|\frac{\sum \text{Failed\_Beat}}{\sum \text{RB}}\right|\right) \times 100\%. \quad (9)$$

The accuracy of step detection was evaluated by calculating the difference (DEF_{SC}) between real step count (RSC) and detected step count (DSC), as shown in (10), using data obtained from the five test subjects at each of the speeds. These values were then compared with those obtained using the two commercial pedometers. The detection rate (DER_{SC}) is defined as follows:

$$\text{DEF}_{\text{SC}} = \text{RSC} - \text{DSC}, \quad (10)$$

$$\text{DER}_{\text{SC}} = \left(1 - \left|\frac{\text{DEF}_{\text{SC}}}{\text{RSC}}\right|\right) \times 100\%. \quad (11)$$

3. Results

3.1. ECG Raw Data. Evaluating the accuracy of the devices began with a comparison of the raw data obtained from the proposed device and that from *CardioSoft*. Figures 9 and 10 present the raw ECG data obtained from the two devices at walking and running speeds, respectively.

Table 2: ECG peak detection results in MIT-BIH ST Change Database.

Data no.	RB (beats)	FP	FN	Failed beats	Data no.	RB (beats)	FP	FN	Failed beats
300	2558	2	0	2	314	2121	6	1	7
301	2497	4	3	7	315	3274	1	36	37
302	2113	3	0	3	316	3351	1	1	2
303	3005	2	13	15	317	2776	0	26	26
304	1852	1	0	1	318	3531	3	6	9
306	6527	4	0	4	320	3135	0	7	7
307	2469	10	0	10	321	2115	3	0	3
308	2299	14	2	16	322	1508	0	4	4
309	5149	0	0	0	323	5290	5	0	5
310	2410	0	2	2	324	1740	3	0	3
311	3009	1	0	1	325	1465	1	0	1
312	2340	12	4	16	326	2075	2	1	3
313	2701	2	0	2	327	1270	1	0	1
Total						72580	81	106	187

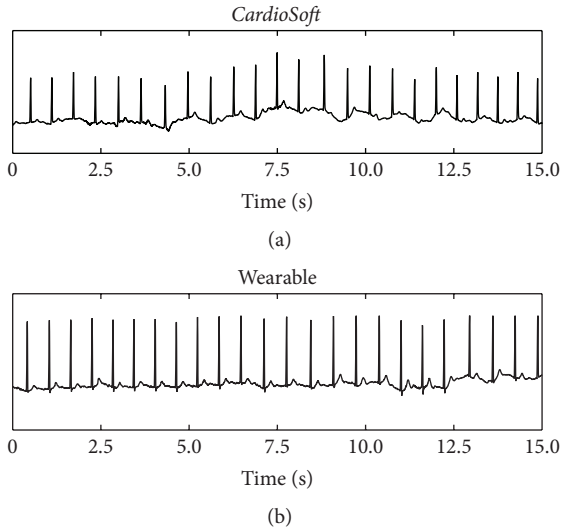


Figure 9: Comparison of raw ECG data under walking conditions: *CardioSoft* (a) and the proposed device (b).

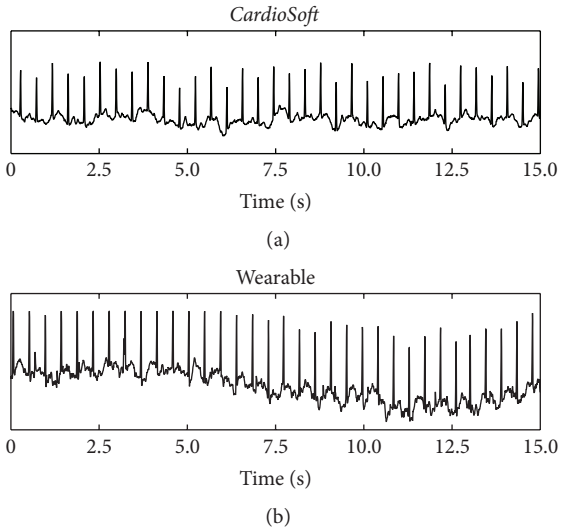


Figure 10: Comparison of raw ECG data under running conditions: *CardioSoft* (a) and the proposed device (b).

3.2. MIT-BIH ST Change Database Test. We applied the proposed ECG peak detection algorithm to 26 datasets from the MIT-BIH ST Change Database (except data numbers 305 and 319), which resulted in a total of 72580 beats. The detection results are listed in Table 2, the sums of which are divided as follows: FP: 81, FN: 106, and total failed beats: 187. Therefore, the accuracy of ECG peak detection based on 10 is 99.7%.

3.3. HR Detection in Treadmill Test. Table 3 lists the HR detection results from one subject in the treadmill test walking at 1.8 km/h. The graph in Figure 11 compares the HR values obtained using the proposed device and those from the *GE Dash 3000* over the period of one minute. In this example, the difference between the HR values from the two devices never exceeds 3 BPM. Table 4 lists the overall results for all

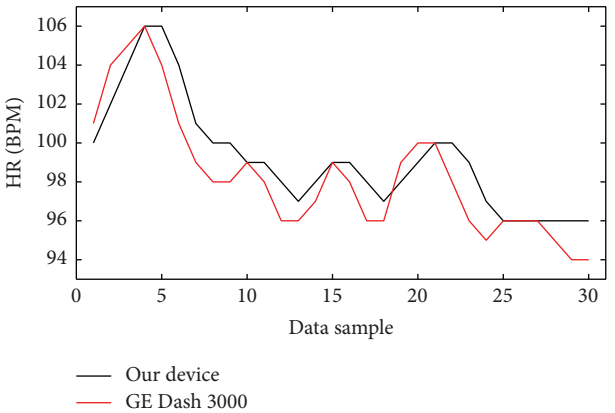


Figure 11: Comparison of HR values obtained from the proposed device and the GE Dash 3000.

Table 3: HR detection accuracy for one subject at 1.8 km/h.

Data no.	Our Device	GE	Err.	Acc. (%)	Data No.	Our Device	GE	Err.	Acc. (%)
1	100	101	−1	99	16	99	98	1	99
2	102	104	−2	98	17	98	96	2	98
3	104	105	−1	99	18	97	96	1	99
4	106	106	0	100	19	98	99	−1	99
5	106	104	2	98	20	99	100	−1	99
6	104	101	3	97	21	100	100	0	100
7	101	99	2	98	22	100	98	2	98
8	100	98	2	98	23	99	96	3	97
9	100	98	2	98	24	97	95	2	98
10	99	99	0	100	25	96	96	0	100
11	99	98	1	99	26	96	96	0	100
12	98	96	2	98	27	96	96	0	100
13	97	96	1	99	28	96	95	1	99
14	98	97	1	99	29	96	94	2	98
15	99	99	0	100	30	96	94	2	98
Average									98.7

Table 4: HR detection results in treadmill test.

Speed (km/h)	1.8	2.7	3.6	4.5	5.4	6.3	7.2	8.1	9.0
Accuracy (%)	98.74	98.67	98.91	98.83	98.92	98.96	98.90	99.06	99.04
Avg. (%)					98.89				

Table 5: Step detection results from one subject in treadmill test.

Treadmill speeds (km/h)	RSC (steps)	A DEF_{SC} (steps)	B DEF_{SC} (steps)	C DEF_{SC} (steps)	A DER_{SC} (%)	B DER_{SC} (%)	C DER_{SC} (%)
1.8	144	−6	−6	−3	95.83	95.83	97.92
2.7	155	0	−3	−26	100.00	98.06	83.23
3.6	164	−2	−2	−16	98.78	98.78	90.24
4.5	173	1	−2	5	99.42	98.84	97.11
5.4	184	0	0	−1	100.00	100.00	99.46
6.3	200	−6	−2	−1	97.00	99.00	99.50
7.2	216	−4	−3	1	98.15	98.61	99.54
8.1	225	1	−10	4	99.56	95.56	98.22
9	237	1	1	9	99.58	99.58	96.20
Average					98.70	98.25	95.71

A: proposed device, B: Omron, and C: Yamax.

five subjects walking or running at all speeds, revealing an overall HR detection accuracy of 98.89%.

3.4. Step Counting Accuracy in Treadmill Test. Table 5 presents the results of step detection for one subject in the treadmill test. The accuracy values of the Yamax SW200 and Omron HJ-720 and the proposed device were 95.71%, 98.25%, and 98.70%, respectively. Table 6 lists the results of step detection for all five subjects walking or running at all speeds, revealing the accuracy of three devices was 92.37%, 94.57%, and 98.96%, respectively.

3.5. Power Consumption. The power consumption of the device was tested in the four main working modes: (1) standby; (2) being not connected; (3) execution of HR/SC algorithm and transmission of calculated data (HR, SC, speed, and distance); and (4) execution of HR/SC algorithm and transmission of calculated and raw (ECG and accelerometer) data. The transmission load for calculated data was two bytes/second, whereas the transmission load for raw ECG data was four bytes/0.01 seconds and raw accelerometer data was two bytes/0.02 seconds. The acquisition of ECG and accelerometer data was, respectively, conducted at sampling

Table 6: Step detection accuracy for all subjects.

Treadmill speeds (km/h)	A DER_{SC} (%)	B DER_{SC} (%)	C DER_{SC} (%)
1.8	98.07	72.04	74.63
2.7	98.19	95.40	76.05
3.6	99.23	98.67	90.56
4.5	99.55	98.94	96.73
5.4	99.32	98.35	98.38
6.3	98.78	98.81	99.04
7.2	98.74	96.08	98.92
8.1	99.22	96.14	99.22
9	99.58	96.74	97.81
Average	98.96	94.57	92.37

Table 7: Current consumption by component in sensor module under various working modes.

Working mode	Sensor module status	Current consumption (mA)
Mode 1	Standby	1
Mode 2	BLE is not connected	6
Mode 3	Execution of HR/SC algorithm and transmission of calculated data	15
Mode 4	Execution of HR/SC algorithm and transmission of calculated and raw data	19

rates of 200 Hz and 50 Hz for each channel. Table 7 lists the consumption of current by the main components in each of the working modes.

4. Discussion

This study developed a wireless wearable system for the monitoring of HR and SC in real time as indications of exercise intensity. The aim was to limit the likelihood of overexertion and/or underexertion in order to maximize the efficiency of training.

4.1. Mobility. The proposed device is compact and lightweight, designed to be attached to the chest of the user over thin clothes. The proposed system monitors 3-axis accelerometer signals and ECG signals without the need to connect electrodes to the subject directly. The proposed system also facilitates continuous remote monitoring via a smartphone using systems such as Mobile-Cloud Telemonitoring [31].

4.2. Accuracy. The detection accuracy of ECG peaks and HR and the number of steps were verified using the MIT-BIH ST Change Database and a treadmill experiment involving five subjects. The average accuracy in ECG peak detection was 99.7%. Detection failures were due to instabilities in the raw ECG data. The accuracy of HR detection remained consistent at 98.89% across all speeds.

The accuracy of step detection was determined using a treadmill test. In most situations, lower treadmill speeds are less accurate than higher treadmill speeds due to the associated signal amplitudes. The Yamax SW200 achieved good performance at walking speeds exceeding 3.6 km/h; however, the Omron HJ-720 was more accurate at lower walking speeds. The proposed system achieved an average accuracy of 98.96%, exceeding that of both commercial devices, regardless of speed.

4.3. Usability and Practicality. Pedometers are a powerful tool for motivating individuals to increase their physical activity; however, unless they are lightweight and easy to use, they will not be adopted. The proposed device overcomes this issue and provides the ability to monitor HR as well as SC.

Furthermore, the power consumption of the device under full load is only 19 mA. The sampling rate could be reduced in order to reduce power consumption even further.

5. Conclusion

This paper outlines an integrated wearable system for the monitoring of HR and SC. Novel step counter algorithm has been developed to achieve the highly accurate step detection. We integrated an accelerometer-based pedometer with a two-electrode ECG circuit to reduce the size and weight of the device. In a treadmill test, the accuracy of the proposed device was shown to be very high. A Bluetooth interface facilitates connection to smartphones for the display, recording, and transmission of data, thereby enhancing the flexibility and usability of the device for health monitoring applications.

Disclosure

Yuan-Hsiang Lin is IEEE Member.

Competing Interests

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

Acknowledgments

This work was supported in part by a research grant from the Ministry of Science and Technology under Grant MOST 104-3115-E-011-001, Avalue Technology Inc., and the Department of Medical Research of National Taiwan University Hospital.

References

[1] WHO, *World Health Report 2010*, WHO, Geneva, Switzerland, 2010, http://whqlibdoc.who.int/whr/2010/9789241564021_eng.pdf.

[2] T. E. Kottke, P. Puska, J. T. Salonen, J. Tuomilehto, and A. Nissinen, "Projected effects of high-risk versus population-based prevention strategies in coronary heart disease," *American Journal of Epidemiology*, vol. 121, no. 5, pp. 697–704, 1985.

[3] R. P. Tonino, "Effect of physical training on the insulin resistance of aging," *American Journal of Physiology—Endocrinology and Metabolism*, vol. 256, pp. E352–E356, 1989.

[4] C. Krucoff, "Popular, low-cost pedometers: 10,000 steps to a better health," *The Seattle Times*, 1999.

[5] C. Tudor-Locke and D. R. Bassett Jr., "How many steps/day are enough? Preliminary pedometer indices for public health," *Sports Medicine*, vol. 34, no. 1, pp. 1–8, 2004.

[6] T. C. T. Ho and X. Chen, "ExerTrek: a portable handheld exercise monitoring, tracking and recommendation system," in *Proceedings of the 11th IEEE International Conference on e-Health Networking, Applications and Services (Healthcom '09)*, pp. 84–88, Sydney, Australia, December 2009.

[7] Z. Li, "Exercises intensity estimation based on the physical activities healthcare system," in *Proceedings of the IEEE International Conference on Communications and Mobile Computing*, vol. 3, pp. 132–136, 2009.

[8] S. M. Fox and W. L. Haskell, "The exercise stress test: needs for standardization," in *Cardiology: Current Topics and Progress*, M. Eliakim and H. N. Neufeld, Eds., pp. 149–154, Academic Press, New York, NY, USA, 6th edition, 1970.

[9] Y. Sun and X. Yu, "An innovative nonintrusive driver assistance system for vital signal monitoring," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 6, pp. 1932–1939, 2014.

[10] E. Khan, F. Al Hossain, S. Z. Uddin, S. K. Alam, and M. K. Hasan, "A robust heart rate monitoring scheme using photoplethysmographic signals corrupted by intense motion artifacts," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp. 550–562, 2016.

[11] D. D. He, E. S. Winokur, and C. G. Sodini, "An ear-worn vital signs monitor," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 11, pp. 2547–2552, 2015.

[12] P. Gupta and T. Dallas, "Feature selection and activity recognition system using a single triaxial accelerometer," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 6, pp. 1780–1786, 2014.

[13] J. Cheng, X. Chen, and M. Shen, "A framework for daily activity monitoring and fall detection based on surface electromyography and accelerometer signals," *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 1, pp. 38–45, 2013.

[14] U. Ryu, K. Ahn, E. Kim et al., "Adaptive step detection algorithm for wireless smart step counter," in *Proceedings of the International Conference on Information Science and Applications (ICISA '13)*, pp. 1–4, Suwon, Republic of Korea, June 2013.

[15] M. Oehler, V. Ling, K. Melhorn, and M. Schilling, "A multichannel portable ECG system with capacitive sensors," *Physiological Measurement*, vol. 29, no. 7, pp. 783–793, 2008.

[16] T. Matsuda and M. Makikawa, "ECG monitoring of a car driver using capacitively-coupled electrodes," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1315–1318, Vancouver, Canada, August 2008.

[17] H. J. Baek, G. S. Chung, K. K. Kim, and K. S. Park, "A smart health monitoring chair for nonintrusive measurement of biological signals," *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 1, pp. 150–158, 2012.

[18] B. Eilebrecht, A. Schommartz, M. Walter, T. Wartzek, M. Czaplik, and S. Leonhardt, "A capacitive ECG array with visual patient feedback," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology*, pp. 6539–6542, Buenos Aires, Argentina, August-September 2010.

[19] J. Wannenburg and R. Malekian, "Body sensor network for mobile health monitoring, a diagnosis and anticipating system," *IEEE Sensors Journal*, vol. 15, no. 12, pp. 6839–6852, 2015.

[20] C. Seeger, K. Van Laerhoven, and A. Buchmann, "MyHealthAssistant: an event-driven middleware for multiple medical applications on a smartphone-mediated body sensor network," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 2, pp. 752–760, 2015.

[21] Bluetooth low energy, http://www.bluetooth.com/.

[22] Texas Instruments, http://www.ti.com/.

[23] D. Dobrev, T. Neycheva, and N. Mudrov, "Simple two-electrode biosignal amplifier," *Medical and Biological Engineering and Computing*, vol. 43, no. 6, pp. 725–730, 2005.

[24] CC2541 bluetooth module, http://www.ti.com/product/cc2541.

[25] E. A. P. J. Prawiro, C.-C. Hu, Y.-S. Chan, C.-H. Chang, and Y.-H. Lin, "A heart rate detection method for low power exercise intensity monitoring device," in *Proceedings of the IEEE International Symposium on Bioelectronics and Bioinformatics (ISBB '14)*, pp. 1–4, April 2014.

[26] R. Libby, "A simple method for reliable footstep detection on embedded sensor platforms," 2008.

[27] MIT-BIH ST Change Database, http://www.physionet.org/physiobank/database/mitdb.

[28] GE Website, http://www3.gehealthcare.com.

[29] B. Dijkstra, W. Zijlstra, E. Scherder, and Y. Kamsma, "Detection of walking periods and number of steps in older adults and patients with Parkinson's disease: accuracy of a pedometer and an accelerometry-based method," *Age and Ageing*, vol. 37, no. 4, pp. 436–441, 2008.

[30] M. L. Moy, A. W. Janney, H. Q. Nguyen et al., "Use of pedometer and internet-mediated walking program in patients with chronic obstructive pulmonary disease," *Journal of Rehabilitation Research & Development*, vol. 47, no. 5, pp. 485–496, 2010.

[31] X. Wang, Q. Gui, B. Liu, Z. Jin, and Y. Chen, "Enabling smart personalized healthcare: a hybrid mobile-cloud approach for ECG telemonitoring," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 3, pp. 739–745, 2014.

Hindawi Publishing Corporation
Mobile Information Systems
Volume 2016, Article ID 7805217, 12 pages
http://dx.doi.org/10.1155/2016/7805217

Research Article

Automatic Gender Detection Based on Characteristics of Vocal Folds for Mobile Healthcare System

Musaed Alhussein,[1] Zulfiqar Ali,[1] Muhammad Imran,[2] and Wadood Abdul[1]

[1]*Department of Computer Engineering, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia*
[2]*Department of Computer Science, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia*

Correspondence should be addressed to Zulfiqar Ali; zuali@ksu.edu.sa

Received 31 December 2015; Accepted 3 April 2016

Academic Editor: Mehmet Orgun

An automatic gender detection may be useful in some cases of a mobile healthcare system. For example, there are some pathologies, such as vocal fold cyst, which mainly occur in female patients. If there is an automatic method for gender detection embedded into the system, it is easy for a healthcare professional to assess and prescribe appropriate medication to the patient. In human voice production system, contribution of the vocal folds is very vital. The length of the vocal folds is gender dependent; a male speaker has longer vocal folds than a female speaker. Due to longer vocal folds, the voice of a male becomes heavy and, therefore, contains more voice intensity. Based on this idea, a new type of time domain acoustic feature for automatic gender detection system is proposed in this paper. The proposed feature measures the voice intensity by calculating the area under the modified voice contour to make the differentiation between males and females. Two different databases are used to show that the proposed feature is independent of text, spoken language, dialect region, recording system, and environment. The obtained results for clean and noisy speech are 98.27% and 96.55%, respectively.

1. Introduction

The applications of automatic gender detection (AGD) system have increased significantly due to the recent developments in speech/speaker recognition, human-computer interaction, and biometric security systems including authentication to access data, surveillance, and security. Gender detection systems limit the search of an imposter to half of the space in many recognition and security systems, where the ultimate goal is the identification of a person. Considering different feature extraction and modeling techniques, an AGD for recognition and security systems should be implemented in such a way that it should not increase the complexity of the whole system. Moreover, a gender detection system can be used for automatic transfer of a phone call of a male/female to the relevant person or department. Furthermore, the accuracy of gender dependent models is higher than gender independent models [1].

In a mobile healthcare system [2–5], automatic gender detection can play a significant role. There are some vocal folds pathologies [6, 7], which are biased to a particular gender; for example, vocal folds cyst can be seen particularly in female patients [8, 9]. If there is a mechanism to automatically detect the gender of the patient, it is easier for a care giver or a healthcare professional to prescribe the appropriate treatment. In this system, the voice or speech of the patient is recorded via a smart device, which is connected to the Internet. The voice or speech is then transmitted to a cloud, where a cloud manager authenticates the patient. The manager distributes the task of feature extraction and classification to various servers, where a decision of gender is made. The decision along with medical data is transmitted to registered healthcare professionals for proper treatment.

In most of the studies [10–16], the acoustic features used for the gender detection depend on the accurate estimation of the fundamental frequency. The accurate estimation of the

fundamental frequency is itself a challenging task. Inaccurate estimation of the fundamental frequency may lead to a significant reduction in the accuracy of a gender detection system. Moreover, various traditional speech features such as linear predictive coefficients (LPC), linear predictive cepstral coefficients (LPCC), Mel-frequency cepstral coefficients (MFCC), perceptual linear predictive coefficients (PLP), and relative spectral PLP coefficients (RASTA-PLP) are used in [10, 12–14, 17, 18] for gender detection. The author in the study [19] claimed that the features used for speech recognition may not provide good results for gender detection. Therefore, it is necessary to explore new features for the gender detection other than traditional speech features, and those features should not rely on the accurate estimation of the fundamental frequency.

In this paper, we proposed a new type of feature for automatic gender detection. The proposed feature considers a speech signal of male and female speakers in time domain and provides a single value in the form of area under the modified voice contour (MVC). The proposed feature does not depend on the estimation of the fundamental frequency, and it has provided good results as compared to the existing features.

Automatic gender detection system based on different types of feature and classifier with varying accuracies are reported in the literature. The acoustic characteristics of humans are based on gender due to physiological changes in glottis, vocal tract thickness, and length. Therefore, researchers are trying to find out the most discriminative features for gender detection. For example, two acoustic features, pitch and first formant, are extracted by linear predictive analysis to construct a gender detection system in [17]. The first feature relates to voice source and the second to the vocal tract. The pitch and the formant frequencies of females are higher as compared to those of males. Euclidean distance and nearest neighbor based classifier has been implemented to detect genders. In the study of Wu and Childers [10], a number of the acoustic parameters, such as autocorrelation, linear prediction, cepstrum, and reflection, extracted from vowels, and voiced and unvoiced fricatives, performed well for gender detection. The study concludes that, for a given gender, the information is time invariant, phoneme independent, and speaker independent. In [11], an accuracy of 96% is attained when pitch is inputted to Multi-Layer Perceptron (MLP) neural network. Pitch, energy, and 12-dimensional MFCC are fed to Support Vector Machine (SVM), and the performance with the gender detection system is 95% [12]. To perform gender detection using vocal source, different vocal source parameters are extracted, and detection rate of 94.7% for male and 95.5% for female voice is achieved in [20].

A comparison between various cepstral features is provided in [13] when extracted from voiced frames only and from a running speech. The cepstral features, MFCC, LPCC, and PLP, are used with their delta coefficients to perform the experiments under three different conditions. Sigmund [18] used selected MFCC to classify the male and female by using short segments of vowels as well as sentences.

A robust gender detection system is developed by Zeng et al. in [14]. The developed system has been tested for the noisy environment, and dependency of the language is also considered for the evaluation of the system. The obtained accuracy is 95%, which shows the robustness of the developed system against noise. The experiment shows that the system is independent of language as well. Relative spectral PLP features and pitch of the male and female speakers are used for gender detection. Chen et al. [15] proposed a gender detection system for children of two age groups of 8-9 years and 16-17 years. The obtained accuracies are 60% and 94%, respectively, for two age groups. Different acoustic features, source spectral magnitude, cepstral peak prominence, and harmonic-to-noise ratio, are used to implement the system. Sedaaghi [21] conducted a comparative study for the gender detection by using two different databases. Various classifiers and acoustic features are used in [16] for gender classification system, and the best reported accuracy is 95%. A total of 113 features are used in the study and Bayes classifier is used for feature selection. The features are grouped into pitch, formant, spectral, and intensity. K-nearest neighbor, SVM, artificial neural networks, and Gaussian mixture model (GMM) are used as classifier in [16]. An automatic gender detection system for Hindi speech is developed in [19] by using MFCC as features and Euclidean distance as a classification method. The authors have mentioned in this study that the same features can be used for gender and speech recognition. However, use of same features for both recognition systems cannot guarantee good performance.

An initial investigation of the proposed feature was done in [22] when MVC was used to measure the voice intensity of the speech sample to discriminate between the genders. The database used was Arabic digits and manual threshold was used to classify the males and females. To authenticate the performance of the proposed feature, TIMIT database was considered and automatic classification of the genders was done by the SVM in [23].

In this study, we have investigated the proposed feature in many aspects, which makes this study different than [22, 23]. The most important factor is to observe the robustness of the proposed feature against noise. Background/environmental noise in speech based applications can degrade the performance of the developed system and, therefore, cannot be neglected. Hence, a white noise at different sound-to-noise ratio (SNR) levels is added in the speech signal of both genders and, then, performance of the developed system is evaluated. The results with clean speech are also obtained to make a comparison between the results of clean and noisy speech. Moreover, many experiments are performed to show that the proposed feature is independent of the language, text, and recording systems. Two databases are used for this purpose; the first database is in English, and the second is in Arabic. Both databases are recorded by using different recording systems, and spoken text is also different. Furthermore, the results of the proposed feature are compared with the pitch plus RASTA-PLP features which provided the best accuracy of 95% in [14]. The proposed feature provided good accuracy and can be used with speaker recognition and biometric security systems to reduce the system's complexity by splitting the search space into two halves.

The rest of the paper is organized as follows: Section 2 describes our proposed automatic gender detection system.

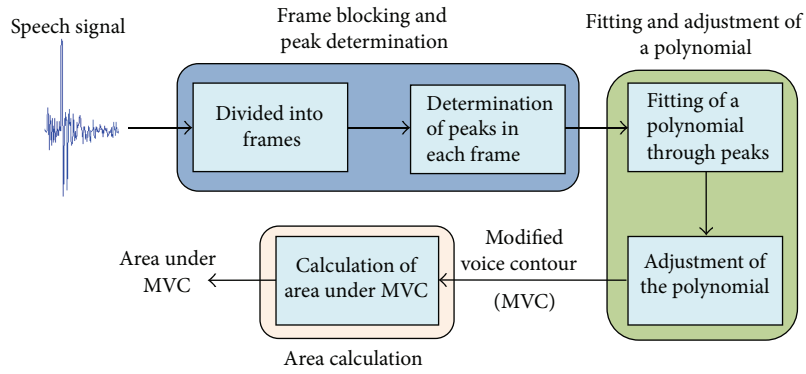


Figure 1: Block diagram to determine modified voice contour and area under it.

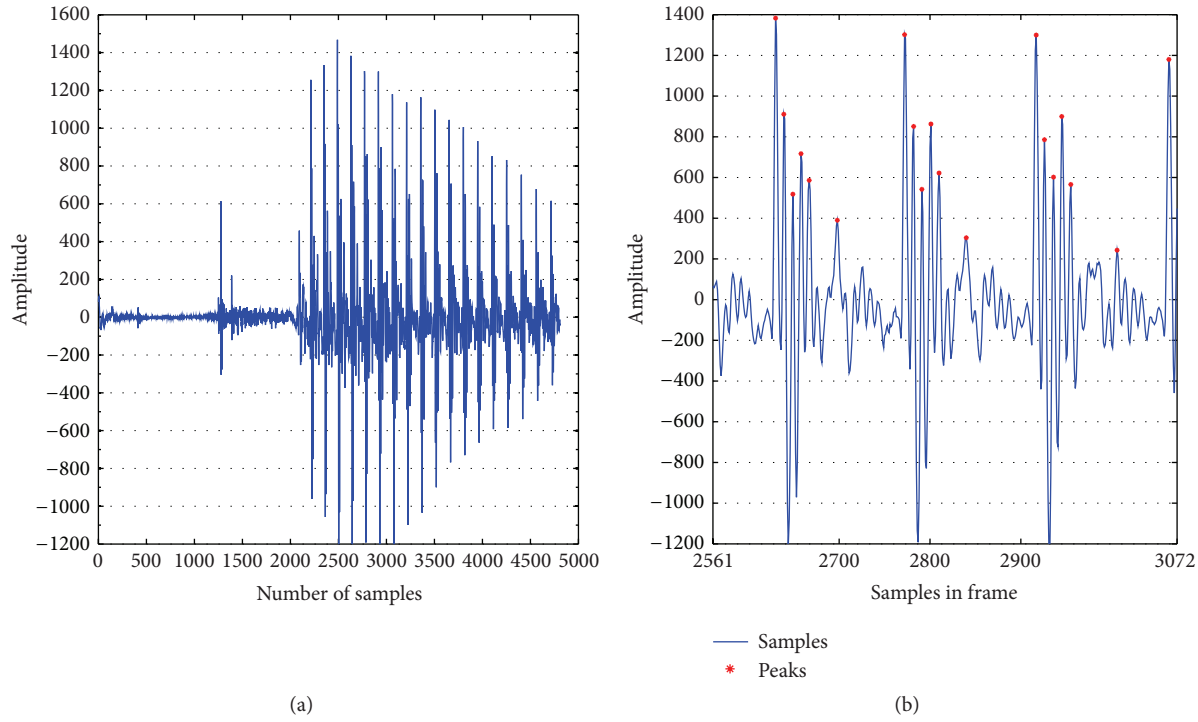


Figure 2: (a) A speech signal. (b) Peaks in a frame.

Section 3 provides the description of speech databases. Section 4 explains the experimental setup and results of the proposed and existing AGD systems. Section 5 analyzes the results and conclusions are drawn in Section 6.

2. Proposed Automatic Gender Detection System

In this study, an automatic gender detection system by using the proposed feature is developed. The proposed feature determines the voice intensity of a speech signal by using the MVC. To implement the feature, Simpson's rule is used to calculate the area under the MVC. The MVC is obtained after adding a factor in a polynomial of degree three that is fitted through the peaks. The peaks are found from each frame when a speech utterance is blocked into frames. At the end, the calculated area is fed to SVM to make the decision about the type of gender. A block diagram to determine the MVC of a speech signal is shown in Figure 1. The implementation of the proposed feature is divided into five major components and they are grouped in three steps: (1) frame blocking and peak calculation, (2) fitting and adjustment of a polynomial, and (3) calculation of area under the MVC by using Simpson's rule. To make a decision about the gender, a binary classifier SVM is used.

2.1. Frame Blocking and Peak Determination. The speech signal, as shown in Figure 2(a), is recorded at the sampling frequency of 16 KHz. A speech signal can be considered dynamic in nature because it changes with time. Therefore,

analysis for whole speech is not possible due to variation in a speech signal. It is the reason that speech may be divided into small frames ranging within 10~40 milliseconds. The variation in size of a frame from 10 to 40 milliseconds is not critical; however, a frame of size 32 milliseconds has provided slightly better results than a frame of lengths 16 and 25 milliseconds [24].

To determine the MVC, peaks are found after blocking the whole speech signal into frames. The length of each frame is 32 milliseconds, and it contains 512 samples. The peaks higher than a certain threshold value thresh are determined in each frame. The thresh is calculated for the whole speech signal by using (1) and kept the same for all the frames of that signal. A frame showing the calculated peaks is depicted in Figure 2(b):

$$\text{thresh} = \left(\text{amp}_{\text{Max}} - \text{amp}_{\text{Min}}\right) * 0.1 + \text{amp}_{\text{Min}}, \quad (1)$$

where amp_{Min} and amp_{Max} are 3% and 97% percentiles of the amplitude in a speech signal, respectively. The thresh varies from signal to signal. Different words exhibit different patterns of the waveform, and, hence, amplitude in a speech signal is also varied. Therefore, to calculate the thresh automatically for each speech signal, (1) is implemented. The relation provided in (1) has also been used successfully in other applications to determine the threshold [25].

2.2. Fitting and Adjustment of the Polynomial. After the calculation of the peaks in each frame, they are joined together. Then, a polynomial of degree three, $g(x)$, is fitted through these peaks to form a curve as shown in Figure 3; the curve is composed of diamonds. It can be observed from Figure 3 that the fitted polynomial passes through the peaks points, hence, not making an envelope over the joined peak points. Therefore, a factor given by (2) is added in the polynomial $g(x)$ to get the MVC:

$$\text{factor} = \left(\max\left(\text{peaks}\right) - \max\left(g\right)\right) * 0.70, \quad (2)$$

where "peaks" is a vector containing peaks of all frames and g is a vector containing all points on the fitted polynomial $g(x)$. Equation (2) provided a factor to adjust the fitted polynomial $g(x)$ which is equivalent to the 70% of the difference between the highest peak and the maximum point on $g(x)$. To avoid the biased peaks, the 70% of the difference is considered; otherwise, the adjusted curve will be misfitted and will not make an envelope over the peaks. After adding the factor in the fitted polynomial, the MVC composed of "$*$" is shown in Figure 3 and is obtained as

$$\text{MVC} = g\left(x\right) + \text{factor}. \quad (3)$$

2.3. Area Calculation. Finally, to obtain the voice intensity, the area under the MVC is calculated with Simpson's rule of numerical integration [26–28], given by (4). The rule divides the region under the MVC into trapeziums, as shown in Figure 4, and calculates area for each trapezium. Then, it takes

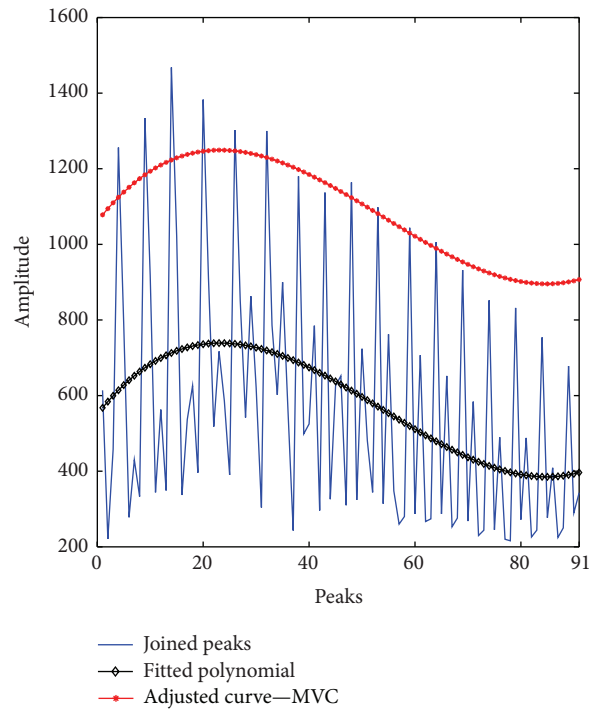


Figure 3: The modified voice contour (MVC).

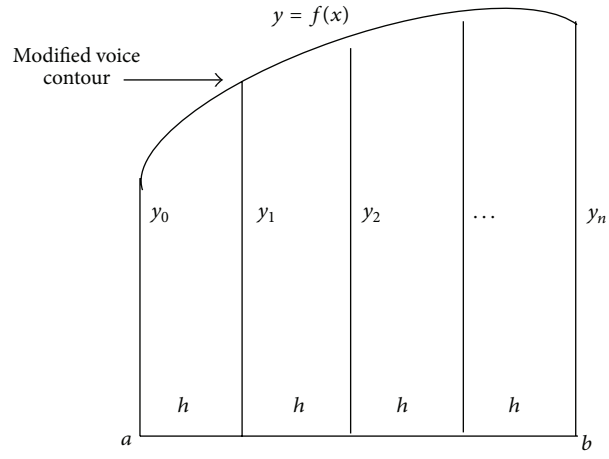


Figure 4: Area under the curve by Simpson's rule.

summation of the area of all trapeziums to provide the total area under the MVC. Simpson's rule calculates area as follows:

$$\text{Area} = \int_a^b f\left(x\right) dx$$

$$\approx \frac{h}{3}\left(y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \cdots + y_n\right), \quad (4)$$

$$h = \left(\frac{b-a}{n}\right),$$

where a and b are the first and the last points on the MVC.

The number of trapeziums under MVC is represented by n. In Simpson's rule, good approximation for area under

a curve can be achieved by increasing n because error in approximation of area decreases as the number of trapeziums increases. The consideration of large value for n is also not feasible because it will increase the computational cost. The value of n is always even in Simpson's rule and it is set to 50 in this study after trying $n = 20, 50, 100$, and 150.

2.4. Support Vector Machine. SVM was proposed by Vapnik [29] and it becomes popular due to its good performance and low computational cost as compared to other classification techniques such as GMM and Hidden Markov Model (HMM). In the developed AGD system, the SVM takes the area under the MVC to make the decision about gender of the speaker. SVM constructs a decision surface (hyper plane) to maximize the distance between two classes, a positive class and a negative class [30]. The dimension of hyper plane depends on the dimension of feature vector given to SVM.

In this study, SVM is implemented by using LIBSVM [24] with a radial basis function (RBF) as kernel, given by

$$K\left(x, x'\right) = \exp\left(-\gamma \left\| x - x' \right\|^2\right), \tag{5}$$

where x is the training sample, x' is the testing sample, and γ is a free parameter. SVM is a linear classifier; however, in most of the cases, data is not linearly separable. Therefore, kernel function is implemented to map the original input space to higher dimensional space, where features are lineally separable. During implementation, male speakers are represented as a positive class and female speakers are represented as a negative class.

3. Material

To evaluate the proposed feature independent of the text and language, two different databases are used. The language of the first database is English, and that of the second is Arabic. Both databases are recorded by using different recording systems and environments, and the spoken text is also different in them.

3.1. TIMIT Database. The DARPA TIMIT Acoustic-Phonetic Continuous Speech Corpus (TIMIT) [31] is used to perform the experiments with English language. The database contains 630 speakers of eight different dialect regions of the United States. Each speaker has recorded 10 sentences at sampling rate of 16 KHz by a condense microphone, where sentence 1 and sentence 2 are the same for all speakers. These fixed sentences are as follows.

Sentence 1. She had your dark suit in greasy wash water all year.

Sentence 2. Do not ask me to carry an oily rag like that.

Only one utterance of these sentences is available because each speaker has recorded these sentences only once. Database includes speakers of many dialect regions but, in this study, we included only those dialect regions in which the total number of speakers is about 100 and contained at least 30 female speakers. Numbers of speakers in dialect regions 2, 4, and 5, labeled as D2, D4, and D5, are 102, 100, and 98, respectively, while the numbers of male and female speakers in each dialect region are listed in Table 1.

Table 1: Number of male and female speakers in different dialect regions.

Dialect regions	Label	Number of speakers		
		Male	Female	Total
2	D2	71	31	102
4	D4	69	31	100
5	D5	62	36	98

Table 2: List of selected words.

Sentence	Word position	Word
1	4	Dark
	5	Suit
	7	Greasy
	8	Wash
	9	Water
	10	All
	11	Year
2	2	Ask
	5	Carry
	7	Oily
	8	Rag
	9	Like

Twelve words, randomly selected, are extracted from sentences 1 and 2. So the total number of available samples for experimentation is 3600 (= 300 × 12). The list of the extracted words is presented in Table 2. The second column provides the position of the word in the sentence. For instance, the first entry of the second column represents that "Dark" is at fourth position in sentence 1.

3.2. The Arabic Database. The Arabic database [32] contains speech samples of 71 speakers: 53 males and 18 females. Each speaker has recorded one utterance of each Arabic digit from one to nine, as listed in Table 3. Speakers recorded the digits with a professional side-address condenser microphone (SHURE PG42) connected to a high-quality mixer (Yamaha MW12CX) at sampling rate of 16 KHz.

4. Experimental Setup and Results

Various experiments are performed to investigate the performance of the proposed feature by using English and Arabic database. Numbers of male and female speakers are not the same in the selected regions D2, D4, and D5 in the English database. So, firstly, one experiment for each dialect region is performed to check the biasness of the proposed feature for gender imbalanced corpus. Secondly, few experiments are performed after making the corpus balanced. Thirdly, noise of different SNRs is added to investigate the robustness of the

Table 3: List of Arabic digits and their IPA.

Digits			
Symbol	In roman English	In Arabic	IPA
1	Wahed	واحد	wa:- ħid
2	Athnayn	أثنين	ʔiθ-ni:n
3	Thalathah	ثلاثه	θa-la:- θah
4	Arbaah	أربعه	ʔar-ba-'ah
5	Khamsah	خمسه	xam-sah
6	Setah	سته	Sit-tah
7	Sabaah	سبعه	Sab-'ah
8	Thamanyah	ثمانية	θa-ma-ni-jah
9	Tesaah	تسعه	tis-'ah

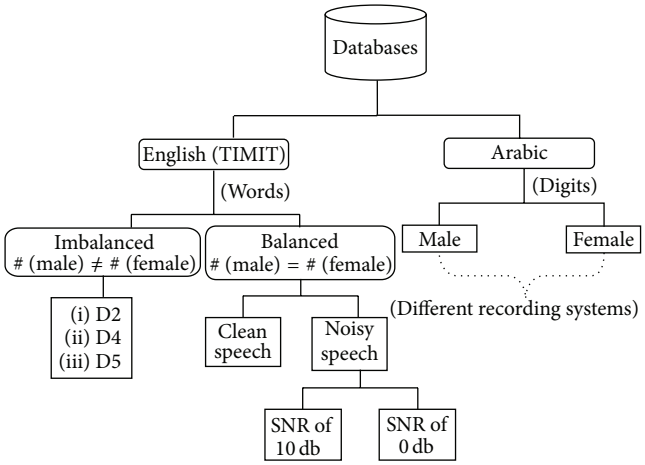


Figure 5: The experimental setup.

proposed feature against noise. Finally, some experiments are performed with Arabic database to observe the accuracy for another spoken language. The text recorded by the speakers in English and Arabic databases is different. The experimental setup is depicted in Figure 5.

All results for gender detection are obtained by using the 5-fold approach. In the 5-fold approach, the database is divided into five distinct subsets. Each time, one of the subsets is used for testing, and the remaining four subsets are used for training of the system. The performance of the proposed AGD system is carried out by using the following parameters:

True positive (TP): the male speaker detected by the system as a male.

True negative (TN): the female speaker detected by the system as a female.

False positive (FP): the female speaker detected by the system as a male.

False negative (FN): the female speaker detected by the system as a male.

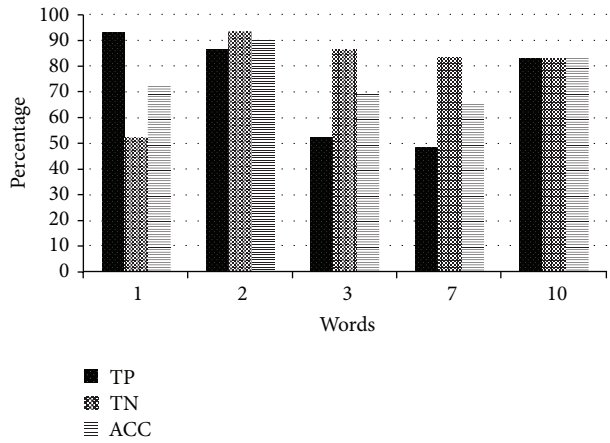


Figure 6: Comparison of performance measures for different words.

Sensitivity (SE): the likelihood that the system detects male when the input is male speaker,

$$\mathrm{SE} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}. \tag{6}$$

Specificity (SP): the likelihood that the system detects female when the input is female speaker,

$$\mathrm{SP} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}}. \tag{7}$$

Accuracy (ACC): the ratio between correctly detected files of the genders and the total number of files,

$$\mathrm{ACC} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \times 100. \tag{8}$$

Area under curve (AUC): the area under the Receiver Operating Characteristic (ROC) curve.

4.1. Gender Detection with Imbalanced Corpus. In this section, the experiments are performed to observe the accuracy of the proposed AGD system by using each word individually and all words simultaneously. In all experiments, numbers of male and female speakers are different.

4.1.1. Gender Detection by Using Individual Words with Imbalanced Corpus. The results of gender detection for each word listed in Table 2 are summarized in Table 4. The dialect region D5 is used in these experiments as it has more number of females as compared to D2 and D4. The highest result for TP is 93% for word 1, and for TN it is also 93% but for word 2. The maximum achieved accuracy is for word 2 and it is 90%. Table 4 provides the analysis of all words in terms of different performance metrics so that we may observe the contribution of each word to gender detection. A comparison of different words having good results in terms of TP, TN, and ACC is depicted in Figure 6.

Table 4: Results for gender detection with individual words by using the proposed feature.

Performance measures	Words											
	1	2	3	4	5	6	7	8	9	10	11	12
TP (%)	93	86	52	72	72	72	48	55	72	83	59	41
TN (%)	52	93	86	48	41	48	83	69	69	83	69	66
FP (%)	48	7	14	52	59	52	17	31	31	17	31	34
FN (%)	7	14	48	28	28	28	52	45	28	17	41	59
ACC (%)	72	90	69	60	57	60	66	62	71	83	64	53

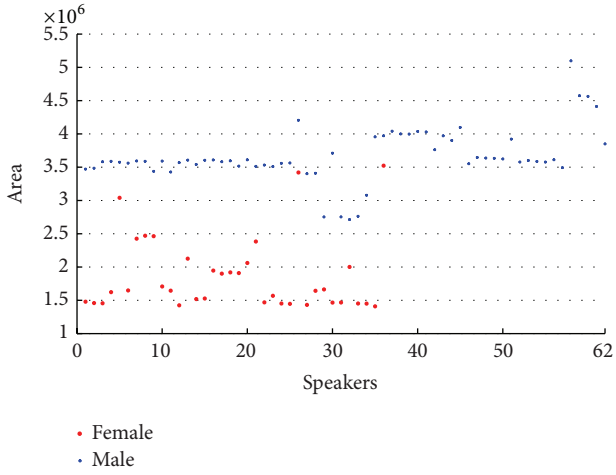


Figure 7: Plot of area under MVC for Word 9 of D5.

4.1.2. Gender Detection by Using All Words Simultaneously with Imbalanced Corpus. Area of all words for each speaker is calculated and, then, fused before providing to SVM for gender detection. Twelve-dimensional feature vector is used for each speaker where each dimension represents an area of the word. The accuracy of 96% is achieved, where obtained SE and SP are 94% and 100%, respectively. Twenty-seven times, out of 28, system detects the gender correctly: 17 out of 18 for males and 10 out of 10 for females. The plot of area under the MVC for male and female speakers for word 9 of dialect region D5 is depicted in Figure 7. Dialect region D5 has 98 speakers: 36 females and 62 males.

By analyzing the results of gender detection by using words individually and all words simultaneously, it can be inferred that it is better to use more than one word to achieve high detection rate. Therefore, in the rest of the experiments, we will use all words simultaneously to achieve better gender detection rate. With all words listed in Table 2, the accuracy of 96% is achieved for D5. The number of male and female speakers in that experiment was 62 and 36, respectively.

It might be assumed that the proposed feature is biased when numbers of samples are different for male and female speakers. To provide the answer about biasness of the proposed feature, it is necessary to use the equal number of male and female speakers for the training and testing of the system. The only available option is to include the female speakers of other dialect regions. It will make a balance between both genders and it will increase the total number of speakers.

Table 5: Results of gender detection by using all words simultaneously for all dialect regions.

Dialect region	Number of speakers (M + F)	Accuracy
D2	71 + 31	94%
D4	69 + 31	96%
D5	62 + 36	96%

M and F stand for male and female speakers, respectively.

However, before doing so, two more experiments for the dialect regions D2 and D4 are performed to investigate the effect of dialect regions. The obtained detection rates for D2 and D4 are 94% and 96%, respectively, which show that dialect regions do not affect the accuracy of the developed system and the performance of the proposed feature is good. Hence, speaker of different dialect regions can be grouped to make the number of male and female speakers equal. The accuracy for D2, D4, and D5 is mentioned in Table 5.

4.2. Gender Detection with Balanced Corpus. It is concluded in Section 4.1 that use of all words simultaneously provided good gender detection rate. Hence, we will continue with it in the rest of the experiments. In addition, the results in Table 5 show that the proposed feature is independent of dialects. Therefore, to make a balance between the numbers of males and females, we can combine speakers of different dialects, and the balanced corpus will be used in the rest of the experiments in this section.

Numbers of female speakers in the dialect regions D2, D4, and D5 are 31, 31, and 36, respectively, and the same numbers of males are taken from these regions to make a balance between male and female speakers. Now, the total number of females is 98 (= 31 + 31 + 36) and the same number of male speakers makes the total number of speakers equal to 196.

A white noise of SNR of 10 db and 0 db is added to the balanced corpus to check the robustness of the proposed feature against noise, and the obtained results are compared with the existing system presented in [14]. The pitch and RASTA-PLP [33] were extracted in [14] from the clean and noisy speech, and eight Gaussians were considered to construct GMM. In this study, to determine the optimized parameters of GMM, mean, covariance matrix, and prior probability, the Expectation-Maximization (EM) algorithm [34] is implemented, while these parameters are initialized by using k-means algorithm [35]. In the GMM based gender detection system, a GMM for both genders is developed. A

Table 6: Results of gender detection for existing and proposed systems with clean speech.

Performance measures	Existing system			Proposed system
	4 Gaussians	8 Gaussians	16 Gaussians	
TP (%)	99.13	99.71	99.71	100.00
TN (%)	91.66	91.66	95.11	96.55
FP (%)	8.33	8.33	4.89	3.45
FN (%)	0.86	0.29	0.29	0.00
SE	0.99	1.00	1.00	1.00
SP	0.92	0.92	0.95	0.97
ACC (%)	95.4	95.68	97.41	98.27
AUC	0.9510	0.9734	0.9795	0.9845

test utterance, for detection of gender of a speaker, will be compared with both models. The model that has a maximum likelihood with the test utterance will be the gender of that test utterance.

4.2.1. Gender Detection for Clean Speech with Balanced Corpus. Two experiments are performed to observe the behavior of the proposed feature for clean speech. In the first experiment, 13 coefficients are extracted per frame in each word. The first coefficient is pitch, and the rest of the twelve features are 11th-order RASTA_PLP coefficients. The extracted features are inputted to the GMM to construct the genders' models by using 4, 8, and 16 Gaussians for male and female detection. The first experiment represents the existing AGD system presented in [14]. This experiment is performed to compare the results with our proposed features.

In the second experiment, the proposed feature provides one value (area under the MVC) for one word which makes the proposed system more efficient. Then, calculated area under the MVC is fed to SVM to make the decision about the gender type. The results of both experiments are provided in Table 6.

The accuracy of the proposed feature is 98.27%, which is better than the existing system. The proposed feature dominates in all performance parameters. The true male detection rate is 100%, and for female, it is 96.55%. Only 3.45% of the female speakers are detected as male, while no male speaker is recognized as female. The experiment for the existing system is performed with the different number of Gaussians to find the best detection rate for that system.

The ROC curve for each system is plotted to analyze their performance, as shown in Figure 8. False positive rate (1 – specificity) and true positive rate (sensitivity) are taken along x-axis and y-axis, respectively. All unique numbers in the decision values of SVM are considered as cut-off points to draw the curve accurately. For existing system, the decision values of the highest accuracy, that is, 97.41%, are used to plot the curve. The area under the ROC curve for the proposed feature is greater than the existing system.

4.2.2. Gender Detection for Noisy Speech with Balanced Corpus. To observe the behavior of the proposed feature in noisy environment, a number of experiments are performed with the speech containing white noise of SNR of 10 db and 0 db, and results are compared with the existing system. The results of the systems for SNR of 10 db and 0 db with different performance parameters are summarized in Tables 7 and 8, respectively.

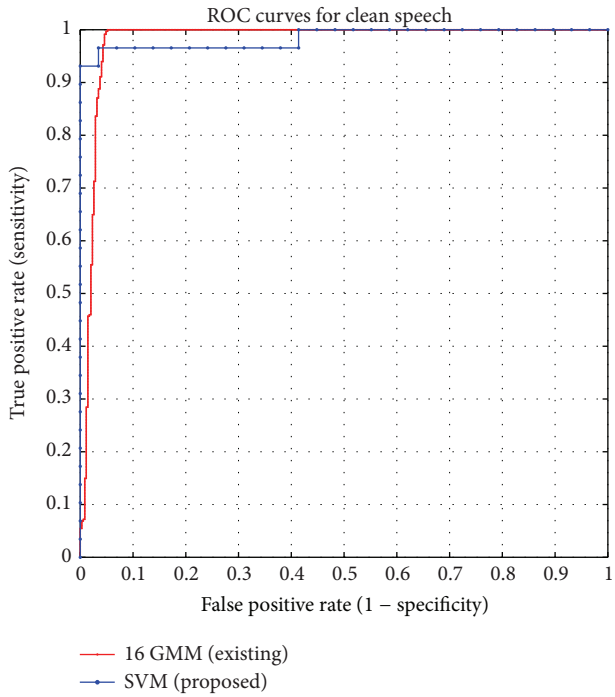


Figure 8: ROC curves of existing and proposed systems for clean speech.

By observing the obtained results, it can be inferred that the feature obtained by calculating area under the MVC can perform well even in noisy environments for the gender detection, and, again, it dominates the RASTA-PLP in most of the performance parameters. The proposed feature has provided accuracies of 96.55% and 94.82% for the SNR of 10 db and 0 db, respectively, which are better than the existing system. The ROC curves for both SNRs for existing and proposed feature are depicted in Figures 9 and 10.

4.3. Gender Detection with Arabic Corpus. To endorse the truths about the proposed feature that it can achieve good

Table 7: Results of gender detection for existing and proposed systems with SNR of 10 db.

Performance measures	Existing system			Proposed system
	4 Gaussians	8 Gaussians	16 Gaussians	
TP (%)	89.37	97.13	96.84	100.00
TN (%)	83.33	89.66	91.09	89.65
FP (%)	16.67	10.34	8.91	10.34
FN (%)	10.63	2.87	3.16	0.00
SE	0.89	0.97	0.97	1.00
SP	0.83	0.90	0.91	0.90
ACC (%)	95.25	93.96	96.12	96.55
AUC	0.9612	0.9732	0.9721	1

Table 8: Results of gender detection for existing and proposed systems with SNR of 0 db.

Performance measures	Existing system			Proposed system
	4 Gaussians	8 Gaussians	16 Gaussians	
TP (%)	89.37	97.13	96.84	100.00
TN (%)	83.33	89.66	91.09	89.65
FP (%)	16.67	10.34	8.91	10.34
FN (%)	10.63	2.87	3.16	0.00
SE	0.89	0.97	0.97	1.00
SP	0.83	0.90	0.91	0.90
ACC (%)	86.35	93.39	93.96	94.82
AUC	0.8867	0.9645	0.9609	0.9893

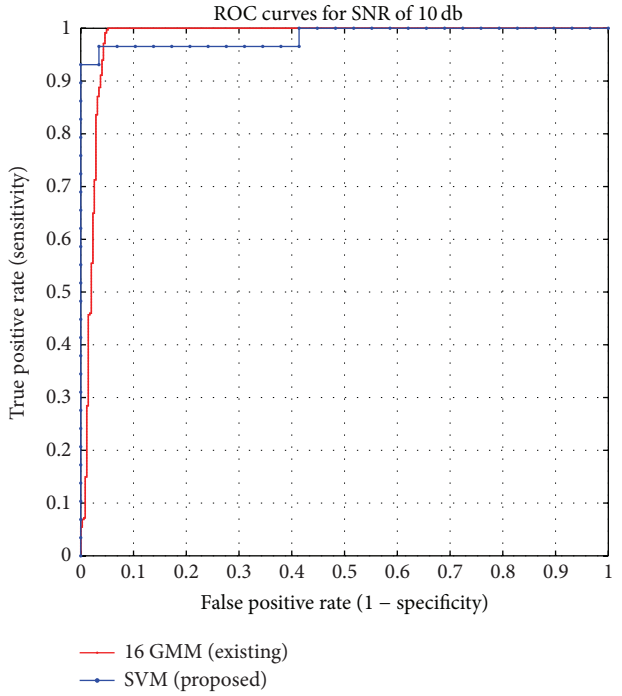


Figure 9: ROC curves of existing and proposed systems for noisy speech 10 db.

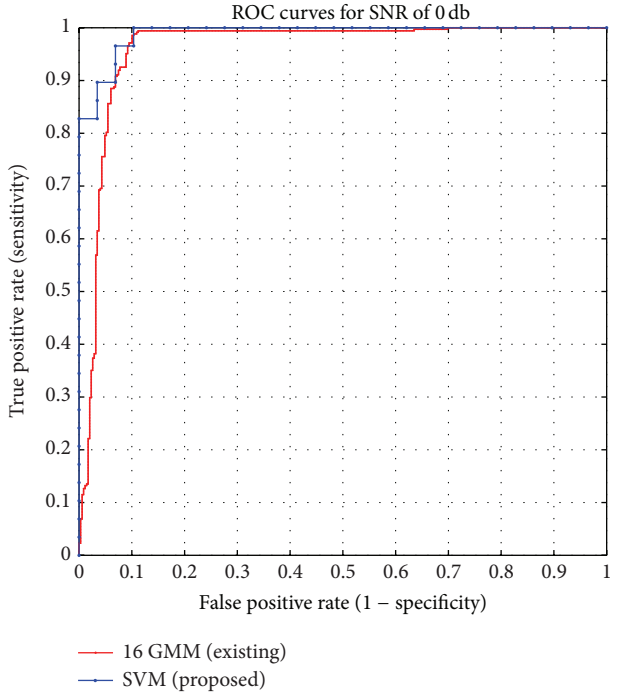


Figure 10: ROC curves of existing and proposed systems for noisy speech of 0 db.

detection rate for other spoken languages and it is independent of text and recording equipment, few experiments are performed. This investigation is performed by using Arabic digits from one to nine, listed in Table 3, uttered by 53 male and 18 female speakers. The area of the MVC is measured by

Table 9: Results of gender detection for existing and proposed systems with Arabic digits.

Performance measures	Existing system			Proposed system
	4 Gaussians	8 Gaussians	16 Gaussians	
TP (%)	92.5	94.3	96.2	100
TN (%)	88.9	94.4	94.4	100
FP (%)	11.1	5.6	5.6	0
FN (%)	7.5	5.7	3.8	0
SE	0.92	0.94	0.96	1.0
SP	0.88	0.94	0.95	1.0
ACC (%)	91.5	94.3	95.7	100
AUC	0.93	0.95	0.97	1.0

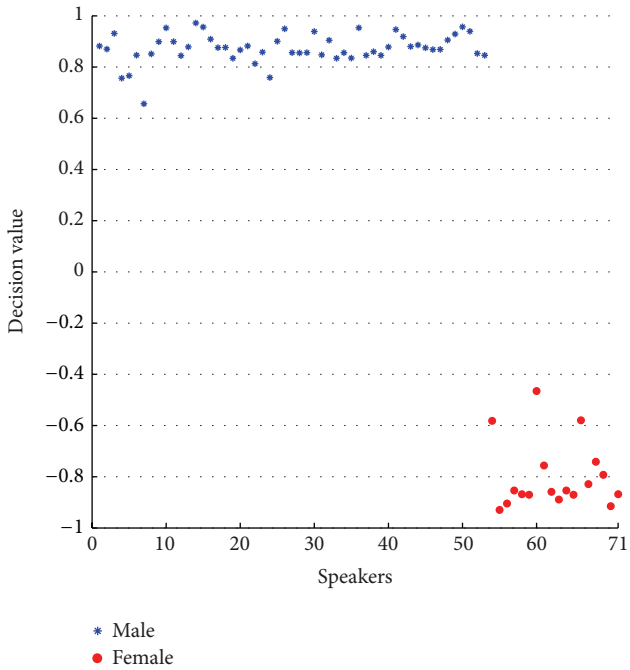


Figure 11: Decision values of the genders for Arabic digits obtained by SVM.

following the steps depicted in Figure 1 and given to SVM for detection of the gender.

Area of each digit is calculated for every speaker, so the dimension of feature vector for each speaker is nine when it is fed to SVM. The decision values obtained, during the classification of genders, by SVM are plotted in Figure 11. The area under the MVC of one or two utterances could be plotted easily, but in this experiment, areas of nine digits are fused. The interpretation of such multidimensional features is not easy to understand by a human mind. Therefore, studies based on multidimensional analysis introduce machine learning stage in the systems [36], and decision values obtained from the classifier can be considered as discriminate measures between classes. Hence, in Figure 11, we have plotted the decision values obtained from SVM.

It can be observed from Figure 11 that the values for the positive and negative classes are perfectively classified. There is no room for the confusion of the genders with each other as they are separated by good margin. The obtained accuracies for the genders are 100%. TP and TN are also 100% while FN and FP are 0% for both male and female. The results of the existing system and proposed system are provided in Table 9. The performance of the proposed system is much better than that of the existing system.

5. Discussion

A noise robust AGD system by using the proposed feature is developed in the study. The proposed feature depends on the voice intensity of a speech signal. In human voice production system, air pressure generated by lungs passes through the windpipe, also known as a trachea. The generated pressure vibrates the vocal folds which reside right on the top of the trachea. The vibration of the vocal folds, open and close, produces a voice which travels through human's mouth and generates speech. The characteristics of the voice vary from person to person due to varying shape, length, and thickness of the vocal folds. Therefore, voices of people feel significantly different from each other. The length of the vocal folds for a human usually lies between 12 and 24 millimeters (mm), whereas the thickness is 3 to 5 mm [37].

The size of the vocal folds also depends on the gender of human beings. The vocal folds length for female is approximately from 12.5 mm to 17.5 mm, while for a male, the length is from 17.5 mm to 24 mm [38]. Due to longer vocal folds, the pitch of male's voice becomes lower, and, therefore, the voice of a male feels heavier than that of a female. The heavy voice contains more voice intensity, and it is the main motivation to propose a new type of feature for gender detection. The proposed feature measures the voice intensity of the speech signal by calculating area under the MVC. It can be seen in Figure 7 that calculated area under MVC for male speakers is larger than that for a female speaker because the voice of a male has more intensity than a female speaker.

The proposed feature does not rely on the accurate estimation of the fundamental frequency which is itself a difficult task. Most of the acoustic features such as formants, harmonic-to-noise ratio, and pitch estimation depend on accurate estimation of the fundamental frequency. If fundamental accuracy is not determined accurately, the systems based on such a type of features may affect the results. In study

[19], the author claimed that traditional speech features may not perform well in a gender detection system. Therefore, a new type of feature is proposed in this study for automatic gender detection.

The developed AGD system has been evaluated in different ways by using clean speech, noisy speech, balanced speech corpus, imbalanced speech corpus, two spoken languages, and different recording systems/environments and text.

For the gender imbalanced corpus, the obtained detection rates are in the range of 94% to 96% for the D2, D4, and D5. After making the gender corpus balance by including the female speakers of the three dialects, the true positives (TP) are 100% and true negatives (TN) are 96.55% for the proposed feature. For the existing system, the best TP and TN are obtained with 16 Gaussians, and they are 99.71% and 96.55%, respectively. The accuracy and the area under the ROC curve for the proposed feature are also better than those for the existing system.

The gender detection rate of the proposed system for noisy speech is also higher than that of the existing system. The TP and TN for both SNRs of 10 db and 0 db are 100% and 89.65%, respectively. For existing system, the obtained TP is 97.13% with 8 Gaussians, and the TN is 91.09% with 16 Gaussians for both SNRs. The area under the ROC curve for the proposed feature is 2.84% better than that for the existing system for 10 db and 3.45% for 0 db.

The proposed feature has provided promising result with the Arabic digits. All males and females are perfectly classified by the feature and obtained detection rate is 100%. The results also show that the feature is capable of detection of genders with any spoken language. Overall, the proposed feature performed well under different circumstances. The word by word experiments for gender detection show that proposed feature can help in finding the words and phonemes that can provide good detection rate.

To observe the statistical significance, Mann-Whitney U test is performed at 5% significant level. The obtained p value for clean speech is $0.10E-5$, for 0 db noise is $0.12E-4$, and for 10 db noise is $0.15E-4$. All p values are less than 0.05 which reject the null hypothesis that decision values of male and female speakers are from continuous distribution with equal medians. The Mann-Whitney U test shows that the proposed system can differentiate between males and females significantly.

6. Conclusion

A new type of feature for the gender detection system is proposed in this study. The developed system can be used in the mobile healthcare systems as it provided good detection rate in both normal and noisy environments. The use of the proposed system with the mobile healthcare systems may assist the doctors in assessing and prescribing appropriate medication to the patient.

The proposed system determines the MVC of the speech signal and then finds area under the MVC to differentiate between male and female speakers. The area under the MVC represents the voice intensity of a speaker. The voice intensity of a speaker is highly dependent on the vocal folds. The size of the vocal folds in a male speaker is longer than that in a female speaker which makes the voice of a male heavy. Therefore, male speakers have more intensity in their voices than females.

Many experiments are performed by using two databases of different languages to evaluate the proposed method and to test its validity under different circumstances. With the help of conducted experiments, we can conclude that the proposed feature can perform equally well in any language. It is unbiased and independent of language, spoken text, and recording equipment. Moreover, the proposed feature is able to provide good detection rate even for the noisy environment. All obtained results are better than the existing AGD system.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

The authors extend their appreciation to the Deanship of Scientific Research at King Saud University, Riyadh, Saudi Arabia, for funding this work through the research group Project no. RG-1436-016.

References

[1] H. Harb and L. Chen, "Voice-based gender identification in multimedia applications," *Journal of Intelligent Information Systems*, vol. 24, no. 2, pp. 179–198, 2005.

[2] M. Shamim Hossain and G. Muhammad, "Cloud-assisted Industrial Internet of Things (IIoT)-enabled framework for health monitoring," *Computer Networks*, 2016.

[3] G. Muhammad, "Automatic speech recognition using interlaced derivative pattern for cloud based healthcare system," *Cluster Computing*, vol. 18, no. 2, pp. 795–802, 2015.

[4] M. Shamim Hossain, G. Muhammad, M. F. Alhamid, B. Song, and K. Al-Mutib, "Audio-visual emotion recognition using big data towards 5G," *Mobile Networks and Applications*, 2016.

[5] M. S. Hossain, "Cloud-supported cyber-physical localization framework for patients monitoring," *IEEE Systems Journal*, 2015.

[6] G. Muhammad, T. A. Mesallam, K. H. Malki, M. Farahat, M. Alsulaiman, and M. Bukhari, "Formant analysis in dysphonic patients and automatic Arabic digit speech recognition," *BioMedical Engineering Online*, vol. 10, article 41, 2011.

[7] G. Muhammad, M. AlSulaiman, A. Mahmood, and Z. Ali, "Automatic voice disorder classification using vowel formants," in *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME '11)*, pp. 1–6, Barcelona, Spain, July 2011.

[8] M. Bouchayer, G. Cornut, E. Witzig, R. Loire, J. B. Roch, and R. W. Bastian, "Epidermoid cysts, sulci, and mucosal bridges of the true vocal cord: a report of 157 cases," *The Laryngoscope*, vol. 95, no. 9, pp. 1087–1094, 1985.

[9] M. M. Johns, "Update on the etiology, diagnosis, and treatment of vocal fold nodules, polyps, and cysts," *Current Opinion in Otolaryngology and Head and Neck Surgery*, vol. 11, no. 6, pp. 456–461, 2003.

[10] K. Wu and D. G. Childers, "Gender recognition from speech. Part I: coarse analysis," *Journal of the Acoustical Society of America*, vol. 90, no. 4 I, pp. 1828–1840, 1991.

[11] S. M. R. Azghadi, M. R. Bonyadi, and H. Sliahhosseini, "Gender classification based on feedforward backpropagation neural network," in *Artificial Intelligence and Innovations 2007: From Theory to Applications: Proceedings of the 4th IFIP International Conference on Artificial Intelligence Applications and Innovations (AIAI 2007)*, C. Boukis, L. Pnevmatikakis, and L. Polymenakos, Eds., vol. 247 of *IFIP The International Federation for Information Processing*, pp. 299–304, Springer, Berlin, Germany, 2007.

[12] S. Gaikwad, B. Gawali, and S. C. Mehrotra, "Gender identification using SVM with combination of MFCC," *Advances in Computational Research*, vol. 4, no. 1, pp. 69–73, 2012.

[13] M. Pronobis and M. Magimai-Doss, "Analysis of F0 and cepstral features for robust automatic gender recognition," Tech. Rep. Idiap-RR-30-2009, Idiap, 2009.

[14] Y.-M. Zeng, Z.-Y. Wu, T. Falk, and W.-Y. Chan, "Robust GMM based gender classification using pitch and RASTA-PLP parameters of speech," in *Proceedings of the International Conference on Machine Learning and Cybernetics*, pp. 3376–3379, Dalian, China, August 2006.

[15] G. Chen, X. Feng, Y. Shue, and A. Alwan, "On using voice source measures in automatic gender classification of children's speech," in *Proceedings of the 11th Annual Conference of the International Speech Communication Association (INTERSPEECH '10)*, pp. 673–676, Chiba, Japan, 2010.

[16] F. Lingenfelser, J. Wagner, T. Vogt, J. Kim, and E. André, "Age and gender classification from speech using decision level fusion and ensemble based techniques," in *Proceedings of the 11th Annual Conference of the International Speech Communication Association (INTERSPEECH '10)*, pp. 2798–2801, Chiba, Japan, September 2010.

[17] K. Rakesh, S. Dutta, and K. Shama, "Gender recognition using speech processing techniques in labview," *International Journal of Advances in Engineering & Technology*, vol. 1, no. 2, pp. 51–63, 2011.

[18] M. Sigmund, "Gender distinction using short segments of speech signal," *International Journal of Computer Science and Network Security*, vol. 8, no. 10, pp. 159–162, 2008.

[19] D. S. Deiv, Gaurav, and M. Bhattacharya, "Automatic gender identification for hindi speech recognition," *International Journal of Computer Applications*, vol. 31, no. 5, pp. 1–8, 2011.

[20] V. N. Sorokin and I. S. Makarov, "Gender recognition from vocal source," *Acoustical Physics*, vol. 54, no. 4, pp. 571–578, 2008.

[21] M. H. Sedaaghi, "A comparative study of gender and age classification in speech signals," *Iranian Journal of Electrical Electronic & Engineering*, vol. 5, no. 1, pp. 1–12, 2009.

[22] M. Alsulaiman, Z. Ali, and G. Muhammad, "Gender classification with voice intensity," in *Proceedings of the 5th European Modeling Symposium of Mathematical Modeling and Computer Simulation*, pp. 205–209, Madrid, Spain, November 2011.

[23] M. Alsulaiman, Z. Ali, and G. Muhammad, "Voice intensity based gender classification by using simpson's rule with SVM," in *Proceedings of the 19th International Conference on Systems, Signals and Image Processing*, pp. 570–573, Vienna, Austria, April 2012.

[24] I. Mporas, T. Ganchev, E. Kotinas, and N. Fakotakis, "Examining the influence of speech frame size and number of cepstral coefficients on the speech recognition performance," in *Proceedings of the 12th International Conference on Speech and Computer*, pp. 1–6, Moscow, Russia, 2007.

[25] R. Jang, Audio Signal Processing and Recognition: End-Point Detection in Time Domain, March 2016, http://mirlab.org/jang/books/audioSignalProcessing/epdTimeDomain.asp?title=6-2%20EPD%20in%20Time%20Domain%20(%BA%DD%C2I%B0%BB%B4%FA%A1G%AE%C9%B0%EC%AA%BA%A4%E8%AAk).

[26] C. F. Gerald and P. O. Wheatley, *Applied Numerical Analysis*, Pearson, 7th edition, 2003.

[27] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover, New York, NY, USA, 9th edition, 1972.

[28] A. Horwitz, "A version of Simpson's rule for multiple integrals," *Journal of Computational and Applied Mathematics*, vol. 134, no. 1-2, pp. 1–11, 2001.

[29] S. Haykin, *Neural Networks a Comprehensive Foundation*, McMaster University, Hamil-ton, Ontario, Canada, 2nd edition, 1998.

[30] S. M. Kamruzzaman, A. N. M. Rezaul Karim, S. Islam, and E. Haque, "Speaker Identification using MFCC-Domain support vector machine," *International Journal of Electrical and Power Engineering*, vol. 1, no. 3, pp. 274–278, 2007.

[31] J. S. Garofolo, L. F. Lamel, W. M. Fisher, J. G. Fiscus, D. S. Pallett, and N. L. Dahlgren, "DARPA TIMIT acoustic-phonetic continous speech corpus CD-ROM," Tech. Rep., NIST, 1993.

[32] M. Alsulaiman, G. Muhammad, M. A. Bencherif, A. Mahmood, Z. Ali, and M. Aljabri, "Building a rich arabic speech database," in *Proceedings of the 5th Asia Modeling Symposium (AMS '11)*, pp. 100–105, Kuala Lumpur, May 2011.

[33] M. A. Anusuya and S. K. Katti, "Front end analysis of speech recognition: a review," *International Journal of Speech Technology*, vol. 14, no. 2, pp. 99–145, 2011.

[34] R. A. Redner and H. F. Walker, "Mixture densities, maximum likelihood and the EM algorithm," *SIAM Review*, vol. 26, no. 2, pp. 195–239, 1984.

[35] A. K. Jain and R. C. Dubes, *Algorithms for Clustering Data*, Prentice-Hall, Upper Saddle River, NJ, USA, 1988.

[36] J. I. Godino-Llorente, R. Fraile, N. Sáenz-Lechón, V. Osma-Ruiz, and P. Gómez-Vilda, "Automatic detection of voice impairments from text-dependent running speech," *Biomedical Signal Processing and Control*, vol. 4, no. 3, pp. 176–182, 2009.

[37] M. S. Hahn, B. A. Teply, M. M. Stevens, S. M. Zeitels, and R. Langer, "Collagen composite hydrogels for vocal fold lamina propria restoration," *Biomaterials*, vol. 27, no. 7, pp. 1104–1109, 2006.

[38] I. R. Titze, *Principles of Voice Production*, Prentice Hall, 1st edition, 1994.

Hindawi Publishing Corporation
Mobile Information Systems
Volume 2016, Article ID 3091280, 14 pages
http://dx.doi.org/10.1155/2016/3091280

Research Article

A Case of Engineering Quality for Mobile Healthcare Applications Using Augmented Personal Software Process Improvement

Shahbaz Ahmed Khan Ghayyur, Daud Awan, and Malik Sikander Hayat Khiyal

Faculty of Computer Sciences, Preston University, Islamabad 44000, Pakistan

Correspondence should be addressed to Shahbaz Ahmed Khan Ghayyur; shahbaz.ahmed@iiu.edu.pk

Received 9 January 2016; Accepted 21 February 2016

Academic Editor: Basit Shahzad

Mobile healthcare systems are currently considered as key research areas in the domain of software engineering. The adoption of modern technologies, for mobile healthcare systems, is a quick option for industry professionals. Software architecture is a key feature that contributes towards a software product, solution, or services. Software architecture helps in better communication, documentation of design decisions, risks identification, basis for reusability, scalability, scheduling, and reduced maintenance cost and lastly it helps to avoid software failures. Hence, in order to solve the abovementioned issues in mobile healthcare, the software architecture is integrated with personal software process. Personal software process has been applied successfully but it is unable to address the issues related to architectural design and evaluation capabilities. Hence, a new technique architecture augmented personal process is presented in order to enhance the quality of the mobile healthcare systems through the use of architectural design with integration of personal software process. The proposed process was validated by case studies. It was found that the proposed process helped in reducing the overall costs and effort. Moreover, an improved architectural design helped in development of high quality mobile healthcare system.

1. Introduction

Mobile health technology aids in practice of medicine and is aided by mobile devices. Google Play and App Store currently host nearly 47,000 mobile applications related to healthcare systems and the hits per day exceed 4 million. Hence, mobile health technology is leaping forward and patients and healthcare professionals have started to reap its benefits. This means that the demand of infrastructure and technology developers is rising day by day. This is only possible when the information is stored and retrieved from the management information systems maintained by hospitals and professionals. For better and effective management, it is necessary to control these heterogeneous working groups and the use of technology may help in better coordination of different activities of these groups.

Information Technology and mobile services are considered as a key feature to improve efficiency of the mobile healthcare systems due to adequate support of the systems. Information systems may help in the improvement of the availability [1] and completeness [2], reduce failures [3], and enhance the orientation of comprehensive documents [4–6]. The modern technology has hypnotized the healthcare providers and its adoption has become a necessity in order to equip hospitals with modern trends [7] in order to incorporate those procedures that may help in effective practices related to the medication. These advancements are the result of wireless communications and network technologies in the previous years [6] and have a significant impact on mobile healthcare (M-health) systems and these systems are also termed healthcare systems for mobile computing or medical sensor and communications technologies [8]. M-health is one of the "biggest technology breakthroughs of our time" [9]. Moreover, "M-health technologies have the potential to change every aspect of the healthcare environment, while delivering better outcomes and substantially lowering costs

and at the same time collecting data about healthcare consumers' health status" [10]. The technologies that are involved in M-health are smart and mobile phones, WiFi, and Bluetooth and are used to transmit data in order to facilitate the health services [11].

The use of mobile devices was also successful in the United States in order to deliver the healthcare services and the situation was the same in Europe [12] and in Asia [13]. In developing countries for primary healthcare the key technologies that are used in the development of M-health systems are based on mobile/wireless ICT [14]. Thus, there is an ever growing demand for healthcare related software and system application development professionals. This research paper focuses on quality of work for individual software developers working on mobile healthcare management systems and how they can improve their own productivity and product quality by utilizing an enhanced personal process. PSP has been successfully utilized and implemented in the leading organizations like DEC, AIS, and HP Corporation [15], Motorola Paging Products Group, Signal Inc., Union Switch, and Advanced Information Services Inc. [16]. Seven competency areas are described in the current version of PSPBOK [17] including (1) Fundamental Knowledge, (2) Basic PSP Concepts, (3) Size Measuring and Estimating, (4) Making and Tracking Project Plans, (5) Planning and Tracking Software Quality, (6) Software Design, and (7) Process Extensions and Customization. However, it lacks in addressing architecture design and evaluation capability.

Software architecture has significant implications for mobile healthcare application development. Hence, in order to manage the complexity related to the development, maintenance, and evolution of a critical software-intensive system, the architectural details must be accurate and traceable during implementation [18]. Software architecture plays an important role in contributing towards a software product, solution, or services. It is like a blueprint or skeleton of a software system that is to be built [19]. Software architecture benefits include a tool for stakeholders communication [20], documentation of design decisions, identification of risks as a result of design decisions, and basis for reuse [20], promotes scalability [21], enables scheduling which saves time, cost of correction, or rework, and most importantly helps avoid software disasters [21]. Therefore, the productivity, performance, and quality of product for individual engineers working on mobile e-health applications can be enhanced by incorporating software architecture support into the personal software process for software development.

This research paper first explores the effects of adopting software architecture methodologies into the personal software process and later proving the effectiveness of modified process named Architectural Augmented Personal Process (AAPP). An AAPP is defined and executed with the help of a mobile healthcare system case study for demonstration of personal improvement which in turn results in better quality software systems. The purpose of this research is to find the impact of architecture with respect to the risk, time, cost, and product quality and explore the PSP based software process improvement (SPI) in light of local industry constraints when developing specialized healthcare related software applications.

This paper investigates an augmented process for personal maturity of developers by implementing a mobile application in the domain of healthcare related to the hospital management and patient interaction system for a local hospital in Pakistan, so better quality systems with better risk management and with low cost and tighter schedule, while promoting simplicity feedback efficiency and adaptability, can be produced by incorporating and adapting software architectural practices. Four distinct case studies on two identical systems have been performed and data is recorded for the mobile e-health applications with the help of two teams both having equal skill set and experience from the same organizations. Both are trained in PSP and then the modified AAPP under study and then the results and yields are analysed for both teams to come up with a clear conclusion backed by solid data. The results are then analysed with experience of current system developers who had created the earlier application for the said system with the help of different tools.

2. The Proposed Solution

In order to answer the research questions, an architecture augmented personal software improvement process is proposed which not only does focus on yield of the programmer but also adds the additional benefits of software architecture incorporation which are early risk identification and management, quality and better communication, and rational management of the design and modification decisions while keeping stakeholders in collaboration. This shall in turn result in better time, cost, and scope management and shall also reduce time to market for e-health systems. This section details the activities that resulted in a PSP based architecture centric methodology AAPP. The method, how to customize the personal software process to meet the new challenges, is clearly stated as an 8-step strategy in the SEICMU PSPBOK. The following is the stepwise explanation of the customized PSP named AAPP in this text.

2.1. Steps for Architecture Augmentation. There are eight steps in architecture augmentation in AAPP and their detailed description is as follows.

Step 1 (determine personal needs and priorities). A process incorporating the SEBOK software architecture characteristics is required to be incorporated in the current personal software process which currently as per knowledge area 6 of PSPBOK supports only the detailed design and does not provide any measures to incorporate the architecture design for software product development. This incorporating is required in order to overcome the identified demotivators.

Step 2 (define process objectives, goals, and quality criteria). The incorporated architecture should be lightweight and in alignment with PSP principles and practices and should also encourage simplicity, communication, and feedback for efficiency of the proposed process. Furthermore, the quality

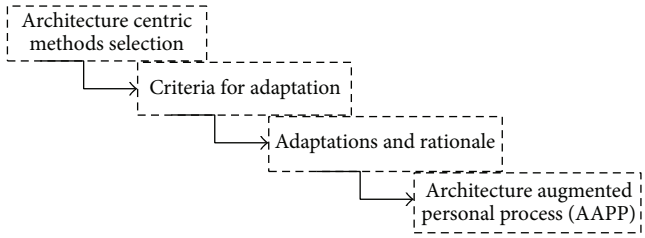


Figure 1: High level activities performed during the proposed methodology development.

and effort along with cost and time to market are not too high when compared to PSP. There are effects on basic measures of time, size, quality (defects), and schedule. Overall impact should also be positive as a result of this augmentation of software architecture into PSP. The proposed process adaptations shall improve the process quality by means of early risk identification and improving overall product quality focusing on quality attributes of interest. The process is also easily adaptable and efficient.

Step 3 (characterize the current process). The current PSP in its present form should be characterized by the PSPBOK.

Step 4 (characterize the target process). Target process should augment the classical PSP with process objectives defined in Step 2. It should also improve quality attributes which are nonfunctional requirements (NFRs). Performance, modifiability, security, and usability are few quality attributes a system may have.

Step 5 (establish a process development strategy). The activities include (1) architecture centric method selection, (2) criteria for adaptation to PP context, (3) adaptations and rationale, and finally (4) Architecture Augmented Personal Process (AAPP). Figure 1 shows key steps involved in the proposed architecture centric method (AAPP) development.

Step 6 (define the initial process). *(a) Architecture Centric Methods Selection*. In our research, SEI QAW [22] and ADD [23] methods have been adapted to form our architecture centric methodology (AAPP) since they are in direct alignment with our research goal for creating an architecture augmented personal process. Software Engineering Institute's methods of quality attribute workshop and attribute driven design are chosen due to their architecture centric nature and origin from the Software Engineering Institute and also for their respective maturity in the area (QAW 3rd Edition, ADD version 2.0) [24]. Moreover, research literature explicitly recommends the integration of architecture centric methods [25] with lightweight software process methodology.

(b) Criteria for Adaptation to Context. To adapt and integrate any process into personal process context, we must acknowledge the ground on which the personal process is based which is taken from capability maturity. Any proposed adaptations must stimulate, have strong grounding in, and be mapped directly with the principles. The criteria have been used for the proposed adaptations and mapped values that stimulate such adaptations, practices that achieve those values, and finally the principles that are key for achieving such values.

Table 1: Quality attribute workshop (QAW) adaptations.

Sr. number	Adaptation to quality attribute workshop
(1)	Presentation and introductions (removed)
(2)	Business/mission presentation (removed)

Table 2: Mapping ADD adaptations to PP values, principles, and practices.

Sr. number	Adaptation to ADD
(1)	More iterative than recursive attribute driven design
(2)	Confirm there is sufficient requirements information (amalgamated)
(3)	Define interfaces for instantiated elements (removed)
(4)	Verify and refine requirements and make them constraints for instantiated elements (amalgamated)
(5)	Repeat from Step 2 to Step 4 (amalgamated with iterative adaptation)

(c) Architecture Centric Methods. SEI quality attribute workshop (QAW) [22] and attribute driven design (ADD) [23] methods were chosen for adaptations according to personal process context. The following subsections map adaptations with personal process values, principles, and practices.

(d) Context to SEI QAW. Table 1 shows the adaptations made to the quality attribute workshop (QAW) [22].

Adapted quality attribute workshop (QAW) consists of six steps. Figure 2 shows the adapted SEI QAW according to AAPP context.

(e) Context Adaptations to SEI ADD. Table 2 shows the attribute driven design (ADD) adaptations mapped to personal process (PP).

Adapted attribute driven design (ADD) according to personal process context consists of four steps that are shown in Figure 3.

(f) Architecture Augmented Personal Process (AAPP) Stage. Figure 4 shows architecture augmented personal process (AAPP) and its activities.

Architecture augmented personal process (AAPP) methodology includes merged activities from adapted quality attribute workshop (QAW) and attribute driven design (ADD) methods as described in [23], respectively. During the second and later iterations for modular development within AAPP, it is made sure that requirements are still consistent with current understanding of the system (Step 1: system stakeholders collaboration) as system is further explored or as a result of communication with stakeholders. Section 3 is about performed case studies in order to verify and validate our proposed AAAP method.

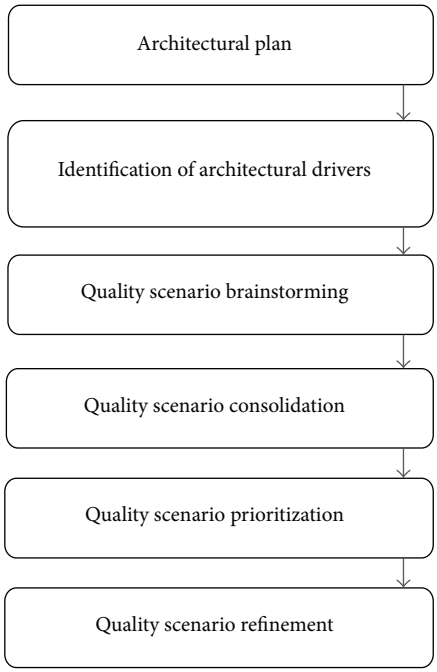


Figure 2: Adapted QAW according to AAPP.

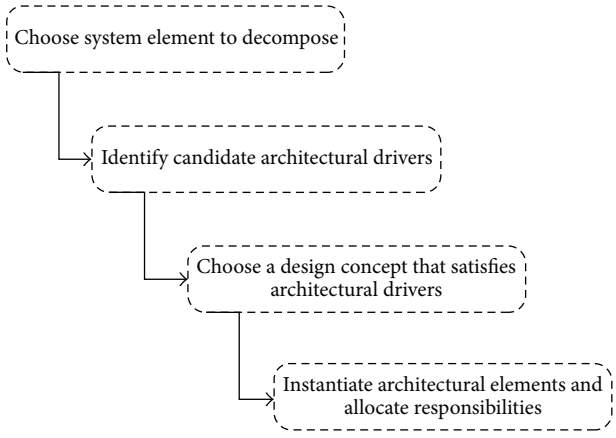


Figure 3: Adapted attribute driven design steps.

3. Case Study

Case study has been chosen as a method to carry out our research since it is a scientific or empirical method used when we want to test whether our theory holds any weight in the real world in a specific context while having no control over variables or context. It is defined as "a technique for detailed exploratory investigations, to understand and explain phenomenon or test theories, using primarily qualitative analysis." The goal of the case study is to provide accurate and comprehensive description of the case. In software engineering, case studies are used for validating research, for example, evaluation of new tools, processes, or methods [25]. Case study is described as a research strategy that comprises design logic, data collection techniques, and specific approaches to data analysis. The following are the strengths and weaknesses of case studies described in [26].

3.1. Case Studies Components. Case studies have the following essential five important components according to Yin [26].

3.1.1. Research Questions. Research questions are usually stated as "who," "what," "where," "how," and "why." Case study is most likely to be used with "how" and "why" questions; however, exceptions and overlaps are common. This research will address the following research questions.

RQ1: How can the software architecture centric methods be effectively incorporated in personal software process?

3.1.2. RQ2. What is the influence of software architecture knowledge support in context to risk, time, cost, and product quality for a process engineer developing mobile applications?

Propositions or Hypothesis. Hypothesis is an educated guess that keeps the research in the right direction [26]. Hypothesis for our research is as follows.

H: Adapted architecture centric method will result in a lightweight quality enhanced approach in terms of effort, cost, risk, and time and quality of product.

3.1.3. Unit of Analysis. There are two methodologies that are compared, that is, personal software process (PSP) and architecture augmented personal process (AAPP) for mobile and e-health application development.

3.1.4. Determination of How Data Are Linked to Propositions. Data collected during case study should be a reflection of the proposition and mapped to it [26].

Validity measures of risk, effort, cost, and quality are key success factors for mobile application software project [27] and they have been used along with time to market and other values' achievement to evaluate the effectiveness of the two approaches. Table 3 shows the validity measures used in this research which reflects our proposition and research questions.

3.1.5. Criteria to Interpret Findings. Any findings and conclusions will be made on the basis of data collected during case studies keeping in view the research questions and propositions along with the statistical analysis.

4. Criteria For Judging Quality of Research Design

Four tests are described in [26] to establish quality of any empirical social research (case study being one of them). These steps are as follows.

4.1. Construct Validity. Construct validity ensures correct operational measures chosen for the concepts being studied. Effort, cost, time to market, and quality attributes achievement levels were measured in this research which reflects

Table 3: Validity measures.

Sr. #	Validity measure	Measured in
(1)	Effort (architecture)	Man-hours
(2)	Effort (overall)	Man-hours for development activity
(3)	Cost (architecture)	Man-hours × wage/hour
(4)	Cost (overall)	Man-hours for development activity × pay/hour
(5)	Time to market (TTM)	TTM = time product delivered − time product conceived
(6)	Quality attributes attainment levels	Percentage (%) of total score achieved

Table 4: Architecture and overall effort measuring matrix.

Day	0-1 hour	1-2 hours	2-3 hours	3-4 hours	4-5 hours	5-6 hours	6-7 hours	7-8 hours
(1)	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D
(2)	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D
N	A/D	A/D	A/D	A/D	A/D	A/D	A/D	A/D

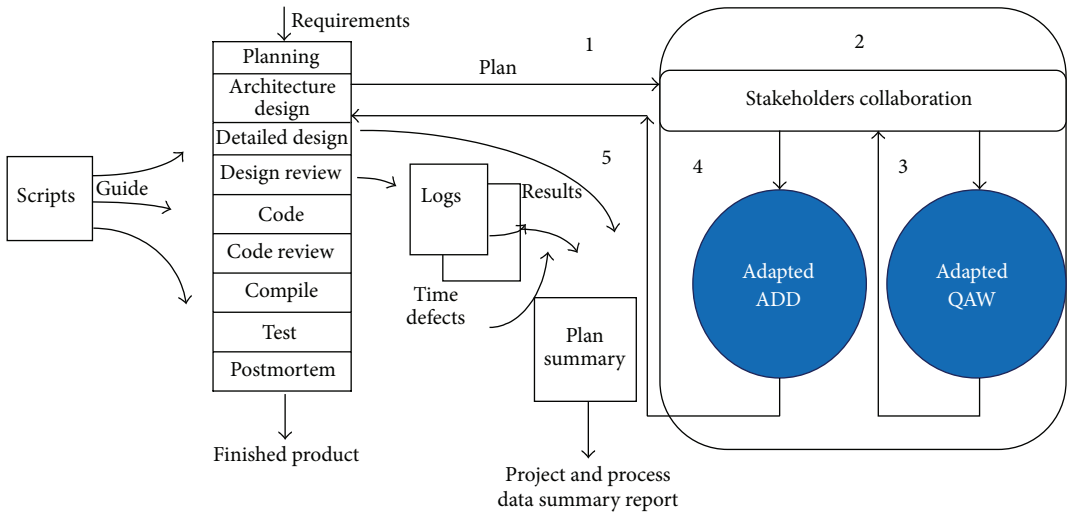


Figure 4: Architecture augmented personal process.

our research questions as well as proposition. These measures have strong architecture literature grounding.

4.2. Internal Validity. Internal validity is inapplicable to case studies that are not concerned with causal situation. In our research, each inference is given its due consideration and rationale during the research design.

4.3. External Validity. Within case studies, it means that the results can be generalized to similar cases to those that were studied. In our research, two separate instances of the same systems were developed to ensure that our results were repeatable and generalizable.

4.4. Reliability. Reliability means that if the same procedures were employed on the same case study again (perhaps by another researcher), the researcher should arrive at the same results/findings earlier recorded. In our research, these steps were documented and executed and are described in Section 7.

5. Measurement Procedures

A day implies 8 working hours while a week is made up of 5 working days. The following were the data measurement procedures for each validity measure.

5.1. Effort. During the day, how much time (man-hours) was spent on architecture activity (marked by an A) or development activity (marked by D) is shown in Table 4.

Effort (architecture) was calculated counting all As and effort (overall) was calculated counting both A and D as shown in Table 4.

5.2. Cost. Architecture cost and overall cost were calculated by multiplying architecture effort and overall effort with wage in $/man-hour, respectively.

5.3. Time to Market. Time to market is the total time taken from project conception to its completion as shown below:

$$\text{Time to Market} = \text{Time Product Delivered} - \text{Time Product Conceived}. \quad (1)$$

5.4. Quality Attributes Achievement Levels. The quality attributes are described along with their subcategories. Each software project has its own software quality attributes of importance in the view of stakeholders. Interviews (client and developers) were used to calculate quality attribute achievement score for each quality attribute; these scores were later aggregated and a holistic percentage (%) score for all quality attributes of importance is calculated.

6. Case Study Execution

The case study was executed in the university lab where two sets of graduate software engineers were selected as participants out of a group of volunteers. Both had 3 years of Management System Development Experience and were also well versed in system design and implementation. Both also have experience in maintaining and documentation of software process data for process improvement. A week of training was also provided to both of the participants of the case study that were provided with adequate training in personal software process and architecture augmented personal process and small examples were executed before moving on to the system under investigation. Identical System Vision documents for alternatively developing the same systems keeping the same lab environment and variables were provided to both sets of engineers along with details of software hardware interfaces available and constraints on the system along with the focused quality attributes of maintainability, usability performance while promoting better communication, simplicity, and feedback for the system.

For comparative analysis and calculation of significance of the effectiveness of both approaches, a small case of mobile application was executed first as a pilot study. The results of the mobile application were analysed to come up with a better understanding of the effectiveness of the proposed process. Time Lapse Assistant redeveloped for this study is an Android OS (or framework) based application and is implemented as part of case studies 1 and 2. Time Lapse Assistant is a tool for professional photographers to calculate time lapse photography parameters and save calculated parameters as part of different projects for later reviewing and other useful tools to aid time lapse photography such as synchronized timer, project map, and sun times.

The main health related mobile application system development undertaken is the hospital management system. This is a term that is mostly referred to where management activities take place residing inside the hospital. The particular system helps the entire hospital management including doctors and clerks. The most important entity that is being focused is the patient. The system is designed to be used on tablets and mobile phones and runs in parallel with the web based system.

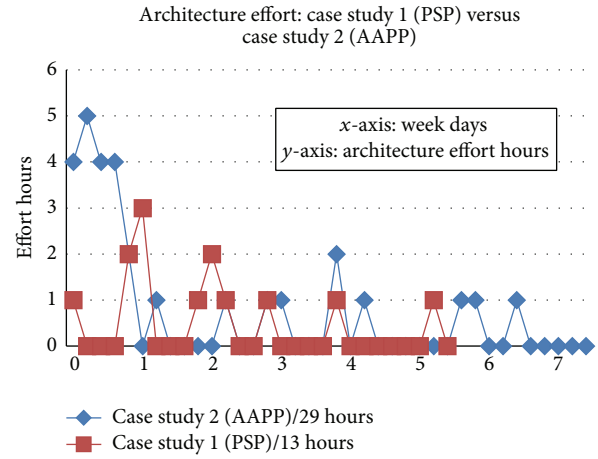


Figure 5: Case study 1 versus case study 2, architecture effort, project size ~2400 LOC.

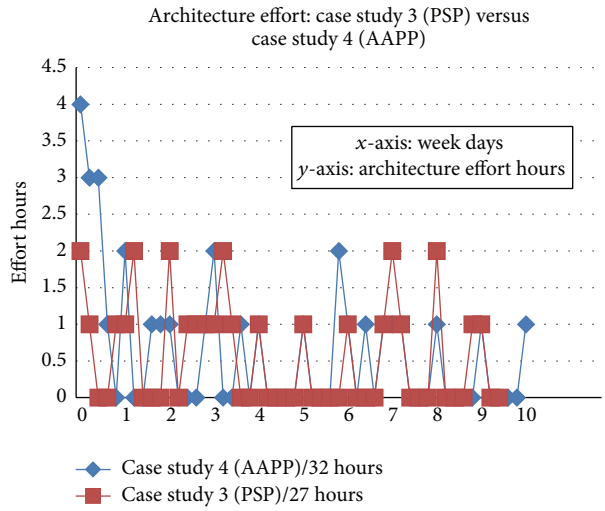


Figure 6: Case study 3 versus case study 4, architecture effort, project size ~4700 LOC.

7. Case Study Results

7.1. Findings: Architecture Effort. Figure 5 presents comparative architecture effort for case study 1, that is, by applying personal software process (PSP), and case study 2, that is, by applying Architecture Augmented Personal Process (AAPP) methodology per week as shown in the figure.

Total architectural effort calculated in *man-hours* for personal software process (PSP) was found to be 13 hours; on the other hand, AAPP took 29 hours (2.23 times greater when compared with PSP) while the project size was approximately 2400 LOC. It was noted that architecture effort for AAPP was considerably higher during the first week due to comprehensive focus on detailed architecture and stakeholder collaboration.

Figure 6 presents comparative architectural effort for case study 3, that is, by applying PSP, and case study 4, by

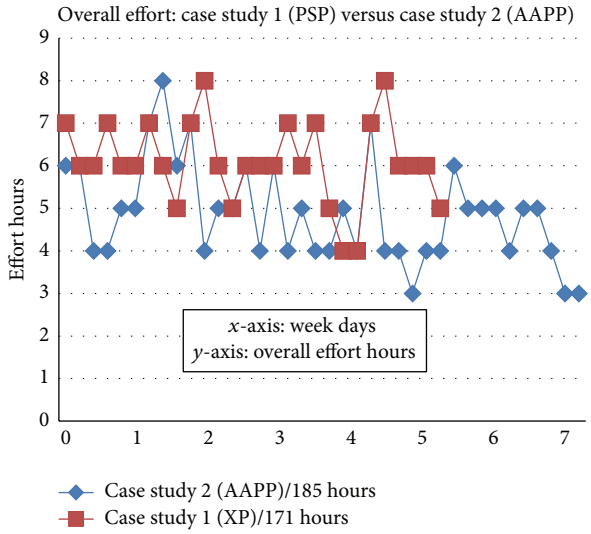


Figure 7: Case study 1 versus case study 2, overall effort.

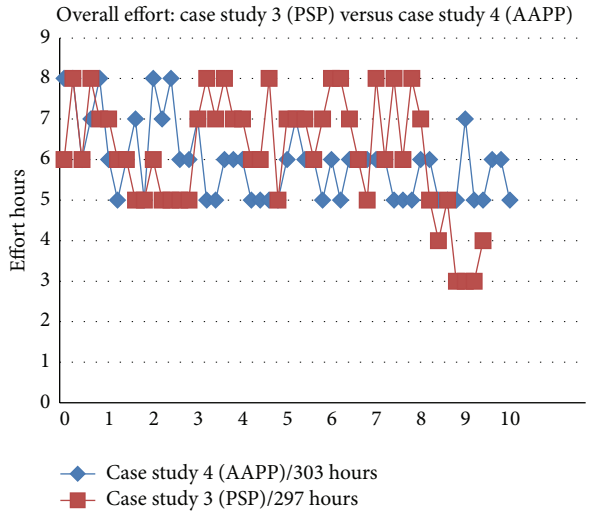


Figure 8: Case study 3 versus case study 4, overall effort.

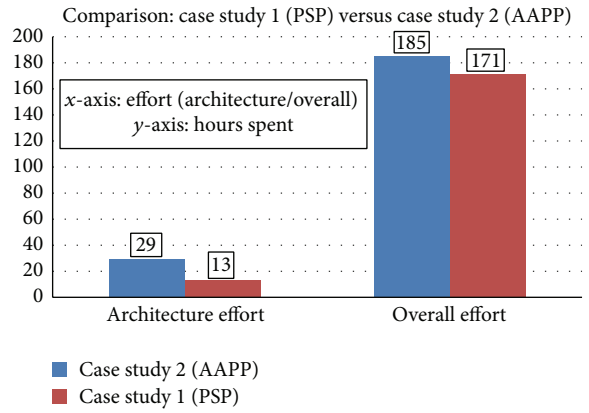


Figure 9: Case study 1 versus case study 2, architecture and overall effort spent.

applying Architecture Augmented Personal Process (AAPP) methodology per week as shown in the figure.

Total architectural effort calculated in *man-hours* for PSP was found to be 27 hours; on the other hand, AAPP took 32 hours (1.18 times greater when compared with PSP design) while the project size was approximately 4700 LOC. Once again, it was noted that design effort for AAPP was higher, but this time not as much significant as with case study 1 and case study 2 comparison, that is, 2.23 times. Furthermore, during case study 3 and case study 4, the project size is greater (4700 LOC) as compared to case study 1 and case study 2; the architecture effort for PSP versus AAPP became less significant, that is, from 16-hour (or 2.23 times greater during AAPP) difference during case study 1 and case study 2 to 5 hours during case study 3/case study 4 (or 1.18 times greater during *AAPP*).

7.2. Findings: Overall Effort. Figure 7 presents overall effort spent (*man-hours*) for the implementation of system during case study 1, that is, by applying PSP, and case study 2, that is, by applying Architecture Augmented Personal Process (AAPP).

During case study 1 (PSP), it took 171 hours, while during case study 2 (AAPP), it took 185 *man-hours* of overall effort. Case study 2 with AAPP took 14 hours more as compared to case study 1. This suggests that investing in architecture effort 2.23 times greater during AAPP saved considerable time during later stages, that is, code, test, and refactoring, and resulted in nonsignificant difference in overall time. Furthermore, there were 4 refactoring activities performed during case study 1 (PSP) suggesting little system understandability and more rework.

Figure 8 presents overall effort spent (*man-hours*) for the implementation of system during case study 3, that is, by applying PSP, and case study 4, that is, by applying architecture augmented personal process (AAPP) as shown in the figure.

During case study 3 (PSP), it took 297 *man-hours* of overall effort while case study 4 (AAPP) took 303 hours. The project sizes for case study 1/case study 2 and case study 3/case study 4 were approximately 2400 LOC and 4700 LOC, respectively. It was found that overall effort or hours got even less significant for AAPP with increased project size, that is, from 14 hours during case study 1/case study 2 to 6 hours during case study 3/case study 4.

7.3. Effort Comparisons for Case Studies. Figure 9 shows architecture as well as overall effort difference for case study 1 and case study 2 between PSP and architecture augmented personal process (AAPP).

Architectural effort in case of AAPP in case study is more than twice but the difference in overall effort is insignificant.

Figure 10 shows architecture as well as overall effort difference for case study 3 and case study 4 between PSP and architecture augmented personal process (AAPP).

With increased project size during case study 3 and case study 4, the difference between both architectural effort and

Table 5: Average architecture and overall effort per day for case studies.

Case study #	Mean architecture effort/day (man-hours/day)	Mean overall effort/day (man-hours/day)
(1) Using PSP	13/28 = 0.4642 hours/day	171/28 = 6.1071 hours/day
(2) Using AAPP	29/38 = 0.7631 hours/day	185/38 = 4.8684 hours/day
(3) Using PSP	27/48 = 0.5625 hours/day	297/49 = 6.0612 hours/day
(4) Using AAPP	32/51 = 0.6274 hours/day	303/51 = 5.9411 hours/day

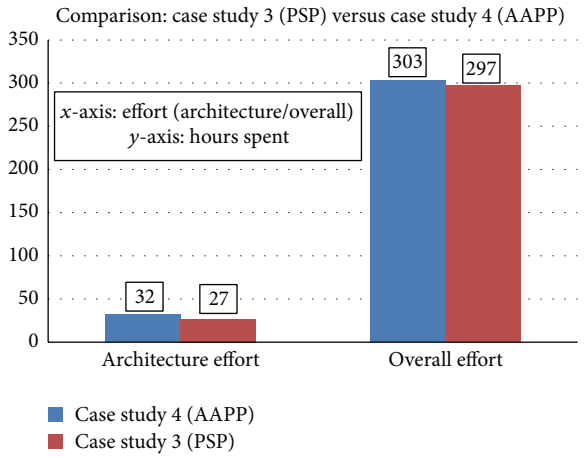


Figure 10: Case study 3 versus case study 4, architecture and overall effort spent.

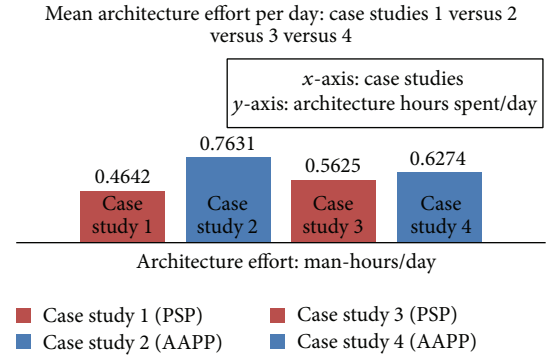


Figure 11: Average architecture effort per day during all case studies.

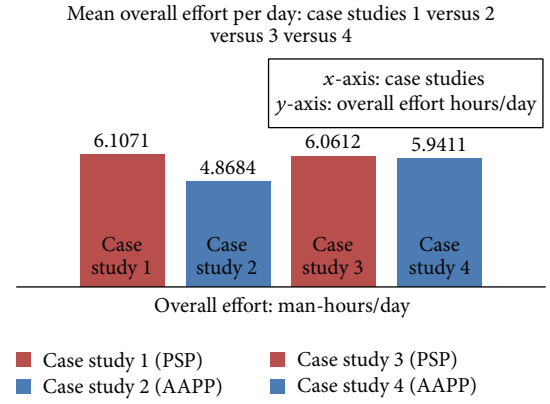


Figure 12: Mean overall effort per day *during* all case studies.

overall effort gets insignificant. This suggests that design or architecture activity for smaller projects may provide its benefits but at doubled architecture effort when architecture augmented personal process (AAPP) was applied compared to traditional architecture activities in personal software process (PSP). However, with increased project size (2400 LOC to 4700 LOC in our case), we notice the effort difference getting less significant between PSP design activities and AAPP. This is because during PSP case studies 1 and 3 more refactoring and rework efforts were made as compared to AAPP case studies 2 and 4 and least planning or architectural effort. Such approach provided good results in terms of less effort spent for PSP architecture for smaller project but with larger project size PSP resulted in even more refactoring and rework effort where the difference between both approaches got even less significant. In other words, the least architectural effort for larger projects resulted in more rework as compared to the smaller projects canceling out any effort saved in the first place.

If we take average for architecture and overall effort, we get the results as shown in Table 5 and Figure 11.

Case study 1 (PSP) resulted in average daily architecture effort of 0.4642 man-hours a day while with case study 2 (AAPP) we get average daily architecture effort of 0.7631 man-hours a day. AAPP requires significantly higher effort as compared to PSP with project of small size. Case study 3 (PSP) resulted in average daily architecture effort of 0.5625 man-hours per day while case study 4 (AAPP) resulted in average daily architecture effort of 0.6274. With an increased project size, we see that average daily architecture effort became insignificant between PSP and AAPP, which is due to higher cost of replanning in case of larger project (case study 3) due to lack of extensive system understanding.

During case studies 1 (PSP) and 2 (AAPP) which represent a small sized project, we get overall daily average effort of 6.1071 and 4.8684 man-hours a day, respectively. This means that with AAPP methodology we saved 1.2387 man-hours on average per day; as more time on architecture was spent during AAPP, this resulted in a better understanding of the system and less rework. However, case studies 3 (PSP) and 4 (AAPP) were conducted on significantly larger project size and the results show an average daily effort of 6.0612 and 5.9411 man-hours, respectively. This difference is insignificant but with AAPP we saved 0.1201 man-hours of overall effort per day. See Figure 12 for further details.

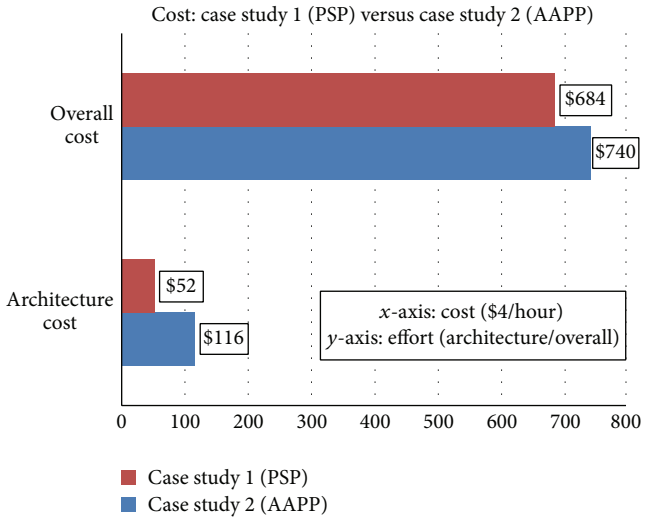


Figure 13: Case study 1 versus case study 2, architecture and overall cost.

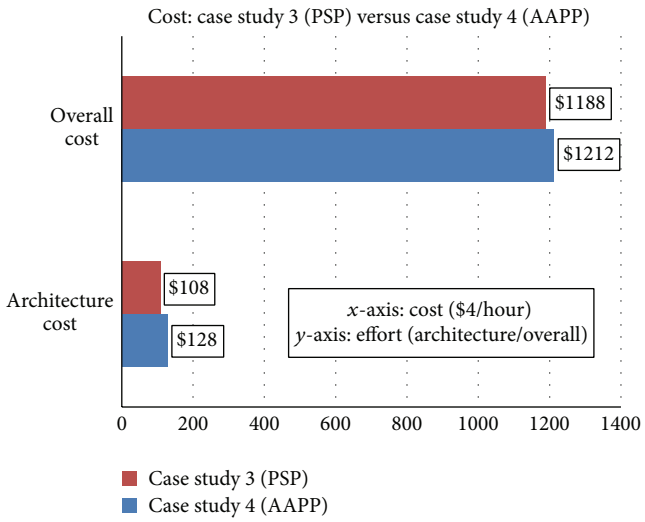


Figure 14: Case study 3 versus case study 4, architecture and overall cost.

8. Findings and Discussions

8.1. Architecture and Overall Costs. Figure 13 shows architecture and overall cost for PSP and architecture augmented personal process (AAPP) approach applied during case study 1 and case study 2, respectively.

During case study 1 and case study 2 with smaller project size (~2400 LOC) it costs more than twice for architecture with AAPP while overall cost is slightly higher as compared to PSP.

Figure 14 shows architecture and overall cost for PSP and architecture augmented personal process (AAPP) approach employed during case study 3 and case study 4, respectively.

As cost is directly proportional to the effort, that is, the more the hours invested in an activity, the greater the cost, AAPP resulted in twice the cost as compared to PSP design during case study 1 and case study 2 but the overall cost difference was not much significant ($66 greater in case of AAPP) as it reduced the amount of rework due to better planning and system understanding that saved time during the later stages (coding, testing, and refactoring). As project size was increased with case study 3 and case study 4, the architecture and overall cost difference between PSP and AAPP became even less significant as compared to case study 1 and case study 2.

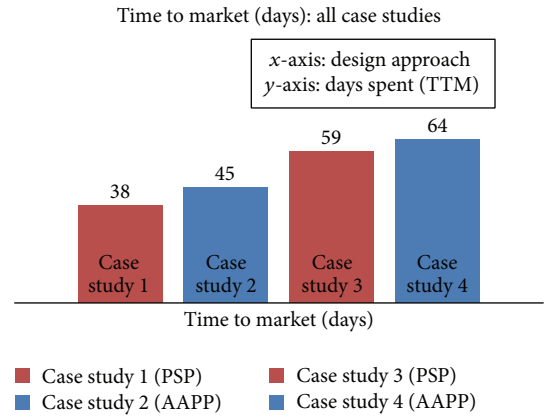


Figure 15: Time to market (TTM).

8.2. Time to Market. Time to market (TTM) is the total time taken for a product as a concept up till when the product is delivered. TTM for case study 1 (PSP Design) was 38 days while for case study 2 (AAPP) it was 45 days. Figure 15 shows TTM for both case studies.

TTM for case study 3 (PSP) was found to be 59 days while TTM for case study 4 (AAPP) was 64 days. The difference between the first and second case study in terms of TTM is 7 additional days in the case of AAPP, that is, case study 2. When compared with overall effort in hours as explained earlier which was 171 and 185 hours for PSP and AAPP, respectively, the difference was found to be 14 hours or less than 2 days of work. Difference between both approaches in terms of overall effort may not be very significant but TTM difference between the two approaches is more significant as compared to overall effort as it also includes the whole week including nonworking days and not just effort spent during working hours. During case study 3 (PSP) and case study 4 (AAPP), TTM difference was found to be 5 additional days for AAPP. Hence, TTM for larger project size (case study 3 and case study 4) was less significant when compared to smaller project size (case study 1 and case study 2) when AAPP was employed.

8.3. Quality Attributes Achievement Levels. Performance/efficiency, modifiability, and usability were found to be common quality attributes for our case studies. Stakeholders were interviewed and data from code was analysed. Each quality attribute achievement score is converted to percentage for ease of understanding.

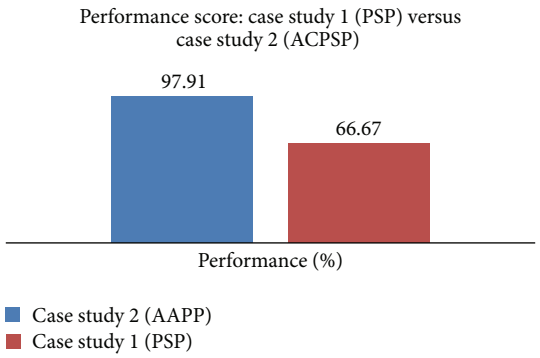


Figure 16: Case study 1 versus case study 2, performance score.

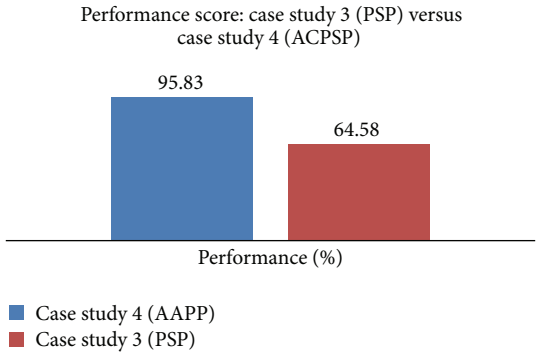


Figure 17: Case study 3 versus case study 4, performance score.

8.3.1. Quality Attribute: Performance Score (%). Performance is the degree to which system or its component meets its required functions within defined constraints, that is, speed, accuracy, or memory usage. *Time economy* is one of the subfactors of performance and is used to measure performance in our case studies as *response time(s)*. Final score is converted into percentage for better understanding. Figure 16 showed the results found for performance quality attribute for case study 1 and case study 2.

Figure 17 showed the results found for performance quality attribute for case study 3 and case study 4.

8.3.2. Quality Attribute: Modifiability Score (%). Modifiability is the ease with which system or component can be modified for corrections of faults, improved performance, adaptations to new environment, or any other attribute. *Cohesiveness*, *correctability*, and *extensibility* subfactors are used to measure modifiability. Final score has been converted into percentage for better understanding. Figure 18 showed the results found for quality attribute modifiability for case study 1 and case study 2.

Figure 19 showed the results found for quality attribute modifiability for case study 3 and case study 4.

8.3.3. Quality Attribute: Usability Score (%). Usability is the ease with which a user can learn to operate, prepare inputs for, and interpret outputs from the system. Subfactors of *learnability* and *satisfaction* are used to measure usability. Final score for usability is converted into percentage for better

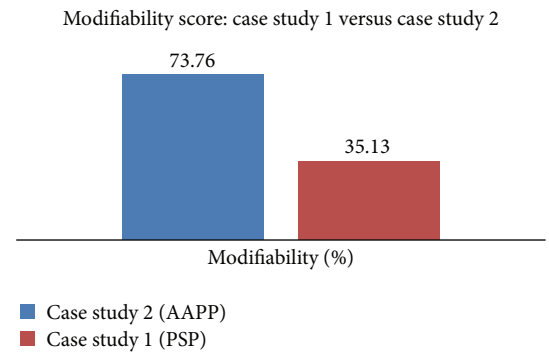


Figure 18: Case study 1 versus case study 2, modifiability score.

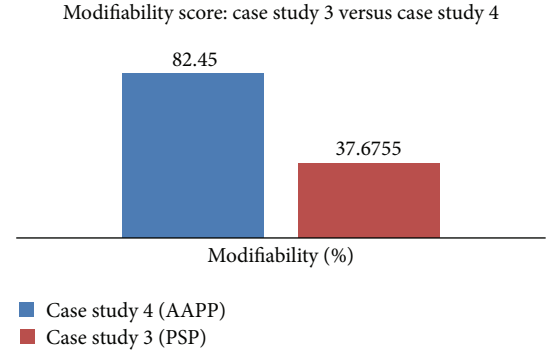


Figure 19: Case study 3 versus case study 4, modifiability score.

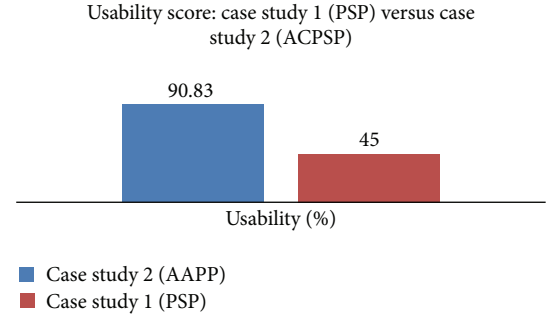


Figure 20: Case study 1 versus case study 2, usability score.

understanding. Figure 20 showed the results found for quality attribute usability for case study 1 and case study 2.

Figure 21 showed the results found for quality attribute usability for case study 3 and case study 4.

8.3.4. Mean Quality Attribute Achievement Score for All Case Studies. Overall results were found for quality attribute achievement levels for case studies 1, 2, 3, and 4 by taking mean of the earlier quality attribute scores in percentages as shown by the following formula:

$$\begin{aligned}\text{Mean QA score for Case Study } X &= \text{QA\% Score 1} + \text{QA\% Score 2} \ldots \\ &+ \frac{\text{QA\% Score } n}{\text{Total QA } (N)},\end{aligned} \tag{2}$$

where QA is the quality attribute.

Table 6: Mean quality attribute scores for all case studies.

Case study #	Mean QA score (%)
(1) PSP	48.93%
(2) AAPP	87.50%
(3) PSP	51.30%
(4) AAPP	91.92%

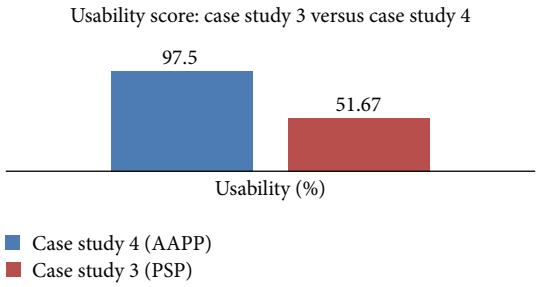


Figure 21: Case study 3 versus case study 4, usability score.

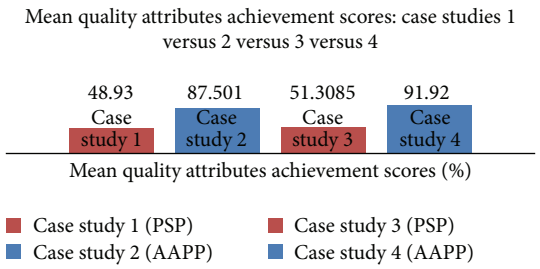


Figure 22: Case studies 1, 2, 3, and 4, mean quality attributes achievement scores.

Simply by adding all quality attributes score within a case study in % and dividing by total number of quality attributes, we get mean quality attribute achievement score for a case study. The mean quality attribute scores are shown in Table 6.

Figure 22 shows mean quality attribute scores for case studies 1, 2, 3, and 4 as shown in the figure.

As shown in Figure 22, the mean quality attribute achievement scores for case studies 2 and 4 by applying AAPP were found to be significantly higher as compared to case studies 1 and 3 by applying PSP.

8.4. Project Values Achievement Score. Project values of simplicity, feedback, and communication were measured by interviewing project stakeholders. First, we presented individual PSP value results and finally looked at the mean PSP values scores achieved during each case study.

8.4.1. Value of Communication Score. Figure 23 shows the results found for the impact on PSP value of *communication* for case study 1 and case study 2 when PSP and AAPP were applied, respectively.

Figure 24 shows the results for the impact on PSP value of *communication* for case study 3 and case study 4 when PSP and AAPP were applied, respectively.

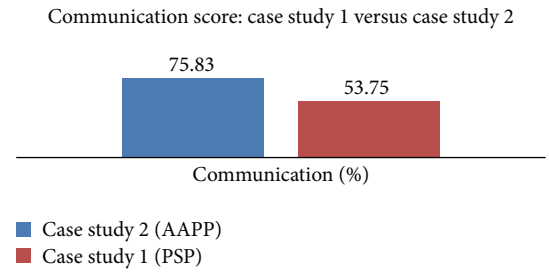


Figure 23: Case study 1 versus case study 2, communication value score.

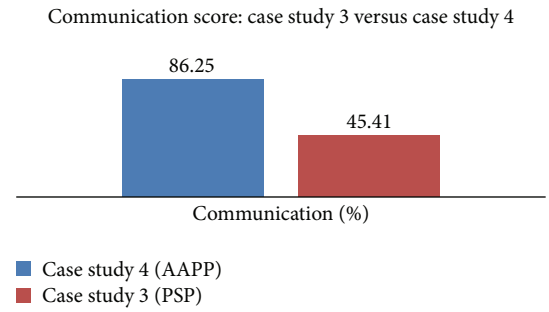


Figure 24: Case study 3 versus case study 4, communication value score.

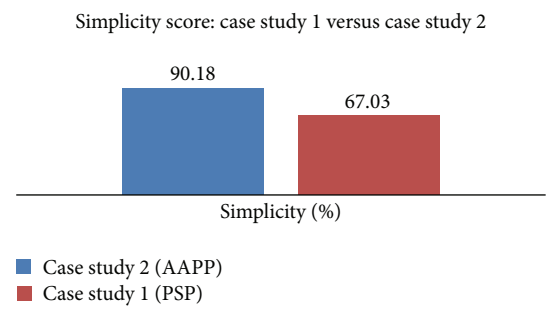


Figure 25: Case study 1 versus case study 2, simplicity value score.

8.4.2. Value of Simplicity Score. These were the results found for the impact on PSP value of *simplicity* for case study 1 and case study 2 when PSP and AAPP were employed, respectively, as shown in Figure 25.

Figure 26 showed the results found for the impact on PSP value of *simplicity* for case study 3 and case study 4 when PSP and AAPP were employed, respectively.

8.4.3. Value of Feedback Score. These were the results found for the impact on PSP value of *feedback* for case study 1 and case study 2 when PSP and AAPP were employed, respectively, as shown in Figure 27.

Figure 28 showed the results found for the impact on PSP value of *feedback* for case study 3 and case study 4 when PSP and AAPP were employed, respectively.

8.5. Architecture Benefits Score. Architecture benefits were measured using interview for personal software process

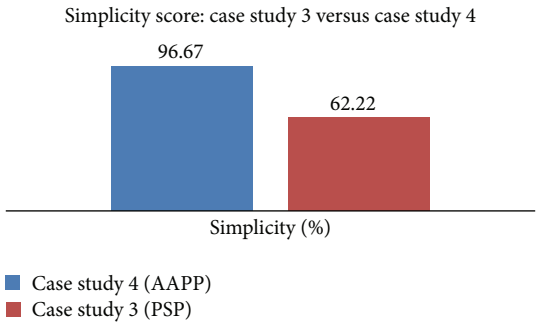


Figure 26: Case study 3 versus case study 4, simplicity value score.

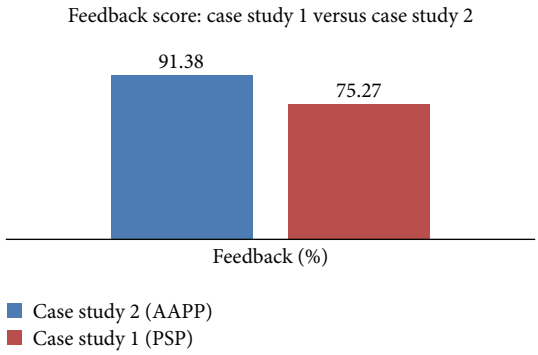


Figure 27: Case study 1 versus case study 2, feedback value score.

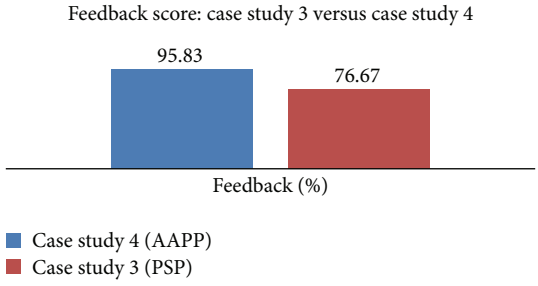


Figure 28: Case study 3 versus case study 4, feedback value score.

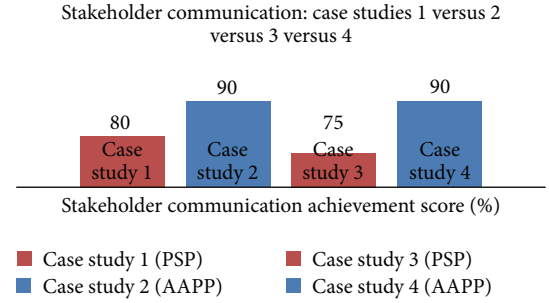


Figure 29: Stakeholder communication score for all case studies.

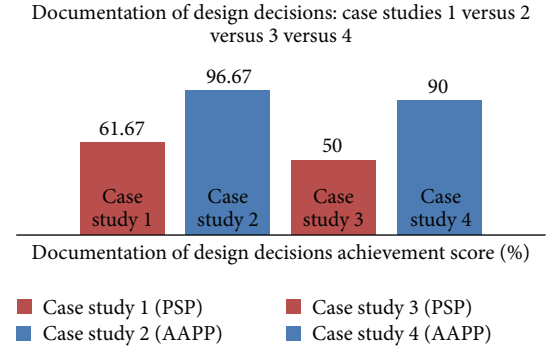


Figure 30: Documentation of design decisions score for all case studies.

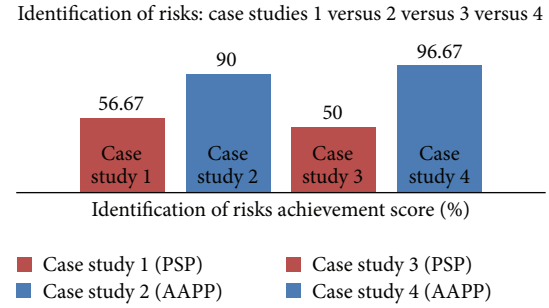


Figure 31: Identification of risks score for all case studies.

(PSP) and architecture augmented personal process (AAPP). Figure 29 shows *stakeholder communication* architecture benefit as percentage score for all case studies.

Figure 30 shows *documentation of design decisions* architecture benefit as percentage score for all case studies.

Figure 31 shows *identification of risks* architecture benefit as percentage score for all case studies.

Figure 32 shows *scalable solutions* architecture benefit as percentage score for all case studies.

Figure 33 shows *scheduling* architecture benefit as percentage score for all case studies.

Figure 34 shows mean architecture benefits score as percentage score for all case studies.

Case studies with architecture augmented personal process (AAPP) showed significantly greater mean achievement of architecture benefits score individually as well as mean, that is, over 20% for project with small size and over 30% for larger sized project as compared to personal software process case studies. It was also noted that team with AAPP had better understanding of the system and as a result created simple solutions that took less time with little or no rework in some cases. This also helped team schedule their tasks better. Table 7 showed the scores recorded for values and quality attributes during case studies 1, 2, 3, and 4; scores are out of 100 and represent percentage (%) as shown in the table.

From the data shown in Table 7, we get a correlation of 0.9923 which means a highly positive correlation and that for its set of data if we increase goal compliance scores, quality attributes scores also increase. In other words, if we stress goal compliance during a project, it has positive impact on quality attributes of that project.

Although the addition of software architecture has induced some additional cost time and effort, the overall advantages outweigh this burden. The results clearly

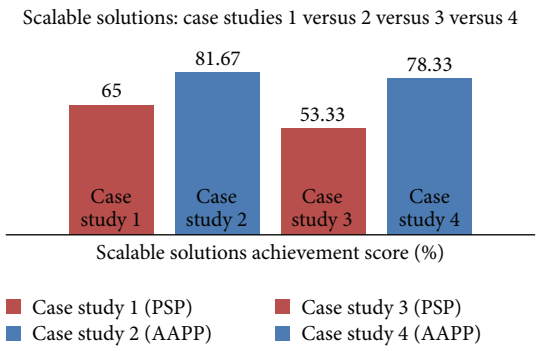


Figure 32: Scalable solutions score for all case studies.

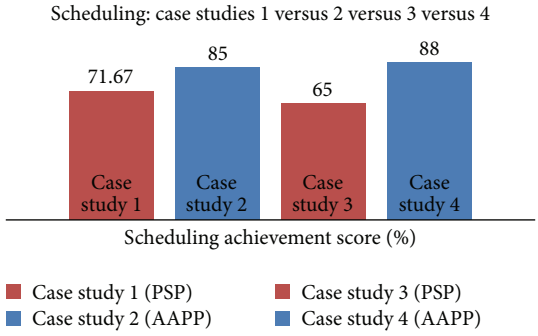


Figure 33: Scheduling score for all case studies.

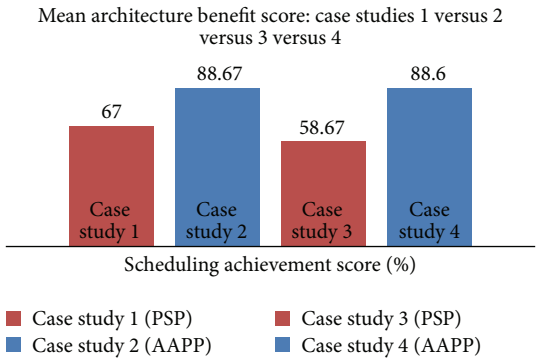


Figure 34: Mean architecture benefit score for all case studies.

show that the augmentation of architecture knowledge has enhanced the overall process output and also impacted the overall product quality. The early risk identification and mitigation also proved to be a handful which reduced the rework time during M-health system development process. Similarly as the size of the project and complexity increases to a certain limit the cost, effort, and time values effect would get reduced for architecture work.

9. Conclusion

The newly proposed process AAPP enhances the overall quality of the mobile healthcare systems in terms of consideration of the architecture with the software process. The case studies were performed in order to compare the proposed process AAPP. The research results show an improvement in terms of architectural benefits that are achieved using AAPP as described in the findings. The application of AAPP is initially a bit costly in terms of time and effort but the performance of the personal process is improved by augmenting the architecture with personal software process. The proposed AAPP is applied in the domain of healthcare in order to get an utmost benefit of the proposed process. Initially, AAPP is applied on a simple healthcare project. In future research, we will apply AAPP on more domains and complex projects in order to better generalize and validate the proposed process.

Table 7: Defined process goals and quality attribute scores for all case studies.

Sr. #	Case study	Process goals score	Quality attributes score
(1)	Case study 1 (PSP)	61.5125	48.93
(2)	Case study 2 (AAPP)	86.325	87.501
(3)	Case study 3 (PSP)	59.5125	51.3085
(4)	Case study 4 (AAPP)	93.125	91.92

Competing Interests

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

Acknowledgments

Special thanks are due to Preston University Pakistan, Participating Engineers, and Higher Education Commission of Pakistan for providing resources in order to complete this research.

References

[1] R. B. Elson, J. G. Faughnan, and D. P. Connelly, "An industrial process view of information delivery to support clinical decision making," *Journal of the American Medical Informatics Association*, vol. 4, no. 4, pp. 266–278, 1997.

[2] P. C. Tang, M. P. Larosa, and S. M. Gorden, "Use of computer-based records, completeness of documentation, and appropriateness of documented clinical decisions," *Journal of the American Medical Informatics Association*, vol. 6, no. 3, pp. 245–251, 1999.

[3] D. W. Bates, A. C. O'neil, D. Boyle et al., "Potential identifiability and preventability of adverse events using information systems," *Journal of the American Medical Informatics Association*, vol. 1, no. 5, pp. 404–411, 1994.

[4] C. J. McDonald and W. M. Tierney, "Computer-stored medical records. Their future role in medical practice," *Journal of the American Medical Association*, vol. 259, no. 23, pp. 3433–3440, 1988.

[5] L. E. Garrett Jr., W. E. Hammond, and W. W. Stead, "The effects of computerized medical records on provider efficiency and quality of care," *Methods of Information in Medicine*, vol. 25, no. 3, pp. 151–157, 1986.

[6] A. Sunyaev, J. M. Leimeister, A. Schweiger, and H. Krcmar, "Integrationsarchitekturen fur das Krankenhaus—Status quo

und Zukunftsperspektiven," *IM-MUNCHEN*, vol. 21, no. 1, pp. 28–35, 2006.

[7] A. Kasbo and R. McLaughlin, *Mobile Health Applications: 2012 Study*, 2012.

[8] R. S. H. Istepanian, E. Jovanov, and Y. T. Zhang, "Introduction to the special section on m-Health: Beyond seamless mobility and global wireless health-care connectivity," *IEEE Transactions on Information Technology in Biomedicine*, vol. 8, no. 4, pp. 405–414, 2004.

[9] S. R. Steinhubl, E. D. Muse, and E. J. Topol, "Can mobile health technologies transform health care?" *The Journal of the American Medical Association*, vol. 310, no. 22, pp. 2395–2396, 2013.

[10] C. E. Kuziemsky, H. Monkman, C. Petersen et al., "Big data in healthcare-defining the digital persona through user contexts from the micro to the macro," *IMIA Yearbook*, vol. 9, no. 1, pp. 82–89, 2014.

[11] F. Imouokhome and V. Osubor, "Mobile-device-based telemedicine for improved health-wealth," *African Journal of Computing & ICT*, vol. 5, 2012.

[12] E. Ammenwerth, A. Buchauer, B. Bludau, and R. Haux, "Mobile information and communication tools in the hospital," *International Journal of Medical Informatics*, vol. 57, no. 1, pp. 21–40, 2000.

[13] S. Hyun, J. Choi, J. Chun et al., "Implementation of mobile computing system in clinical environment: MobileNurse™," in *Proceedings of the AMIA Symposium*, p. 1036, 2000.

[14] R. Wootton, *Telehealth in the Developing World*, IDRC, 2009.

[15] M. Frappier and M. Richard, "SMP: a process-driven approach to project management," in *Proceedings of the 37th Annual Hawaii International Conference on System Sciences*, p. 9, January 2004.

[16] S. B. de Oliveira, R. Valle, and C. F. Mahler, "A comparative analysis of CMMI software project management by Brazilian, Indian and Chinese companies," *Software Quality Journal*, vol. 18, no. 2, pp. 177–194, 2010.

[17] M. Umarji and C. Seaman, "Predicting acceptance of software process improvement," *ACM SIGSOFT Software Engineering Notes*, vol. 30, no. 4, pp. 1–6, 2005.

[18] T. Mens, J. Magee, and B. Rumpe, "Evolving software architecture descriptions of critical systems," *Computer*, vol. 43, no. 5, Article ID 5472890, pp. 42–48, 2010.

[19] N. Rozanski and E. Woods, *Software Systems Architecture: Working with Stakeholders Using Viewpoints and Perspectives*, Addison-Wesley, Boston, Mass, USA, 2011.

[20] A. Rauf, S. Anwar, M. Ramzan, and A. A. Shahid, "Analysis of softfare process improvement efforts in Pakistan," in *Proceedings of the 2nd International Conference on Computer and Automation Engineering (ICCAE '10)*, pp. 375–379, Singapore, February 2010.

[21] D. Garlan and M. Shaw, "An introduction to software architecture," in *Advances in Software Engineering and Knowledge Engineering*, vol. 1, World Scientific Publishing, 1993.

[22] R. Barbacci Mario, "Quality attribute workshops (QAW)," Reporte Técnico del CMU-SEI, 2003.

[23] R. Wojcik, F. Bachmann, L. Bass et al., "Attribute-Driven Design (ADD), Version 2.0," DTIC Document, 2006.

[24] P. Bourque and R. Dupuis, *Software Engineering Body of Knowledge (SWEBOK)*, IEEE Computer Society, EUA, 2004.

[25] S. Easterbrook and J. Aranda, "Case studies for software engineers," in *Proceedings of the International Conference on Software Engineering (ICSE '06)*, Shanghai, China, May 2006.

[26] R. K. Yin, *Case Study Research: Design and Methods*, Sage, Thousand Oaks, Calif, USA, 2013.

[27] R. Atkinson, "Project management: cost, time and quality, two best guesses and a phenomenon, its time to accept other success criteria," *International Journal of Project Management*, vol. 17, no. 6, pp. 337–342, 1999.

Hindawi Publishing Corporation
Mobile Information Systems
Volume 2015, Article ID 260594, 13 pages
http://dx.doi.org/10.1155/2015/260594

Research Article

EVFDT: An Enhanced Very Fast Decision Tree Algorithm for Detecting Distributed Denial of Service Attack in Cloud-Assisted Wireless Body Area Network

Rabia Latif,[1] Haider Abbas,[1,2] Seemab Latif,[1] and Ashraf Masood[1]

[1]*National University of Sciences and Technology, Islamabad 44000, Pakistan*
[2]*King Saud University, Riyadh 11451, Saudi Arabia*

Correspondence should be addressed to Haider Abbas; hsiddiqui@ksu.edu.sa

Received 12 May 2015; Accepted 9 August 2015

Academic Editor: Basit Shahzad

Due to the scattered nature of DDoS attacks and advancement of new technologies such as cloud-assisted WBAN, it becomes challenging to detect malicious activities by relying on conventional security mechanisms. The detection of such attacks demands an adaptive and incremental learning classifier capable of accurate decision making with less computation. Hence, the DDoS attack detection using existing machine learning techniques requires full data set to be stored in the memory and are not appropriate for real-time network traffic. To overcome these shortcomings, Very Fast Decision Tree (VFDT) algorithm has been proposed in the past that can handle high speed streaming data efficiently. Whilst considering the data generated by WBAN sensors, noise is an obvious aspect that severely affects the accuracy and increases false alarms. In this paper, an enhanced VFDT (EVFDT) is proposed to efficiently detect the occurrence of DDoS attack in cloud-assisted WBAN. EVFDT uses an adaptive tie-breaking threshold for node splitting. To resolve the tree size expansion under extreme noise, a lightweight iterative pruning technique is proposed. To analyze the performance of EVFDT, four metrics are evaluated: classification accuracy, tree size, time, and memory. Simulation results show that EVFDT attains significantly high detection accuracy with fewer false alarms.

1. Introduction

Nowadays, cloud-assisted WBAN for patient health monitoring have attracted researchers' attention. Besides other open issues in WBAN environment such as energy efficiency, quality of service, and standardization, security and privacy are the key issues that need special attention. Among these security issues, data availability is the most nagging security issue. The Distributed Denial of Service (DDoS) attack is one of the most powerful attacks on the availability of patients health data and services of health care professional. DDoS attack severely affects the capacity and performance of a WBAN network if not handled in a timely and appropriate manner [1].

For detecting a DDoS attack in cloud-assisted WBAN, there is a need for a defensive approach that understands the network semantics and flow of traffic in the networks. When a victim node is flooded with huge amount of packets that exceeds its processing ability, the excess must be dropped. The packet based dropping strategy helps in distinguishing the legitimate traffic from the flood traffic and is used to avoid the impact of attack traffic on legitimate users. Observing the network traffic flow shows that there is no regular structure of patterns existing in the network and therefore statistical pattern identification techniques are needed. Integrating existing attack detection and defense mechanism in a resource constrained WBAN network increases the computation and communication cost [2].

The network resources are not enough to mitigate the huge amount of traffic generated by DDoS attack. Therefore, there is a need for an approach that is lightweight and capable of handling real-time streaming data. For this research, data mining techniques have been studied and explored. Among the data mining techniques, VFDT is proved to be the most prevalent due to the simplicity and interpretability of their rules and thus considered as more appropriate for

low-power sensor networks. The underlying reasons for the selection of VFDT are as follows: (1) it lightweight; that is, it does not require a dataset to be stored in memory, thus making it suitable for resource constraint WBAN; (2) it can progressively build decision tree from scratch which helps in detecting DDoS attack at any stage; (3) each time a new segment of sensor data arrives, a testing and training process is performed over it keeping the stored data up to date; (4) it does not require reading full dataset and yet adjusts decision tree according to the newly incoming and gathered statistical attributes, thus consuming less memory space; (5) it is appropriate for huge amount of nonstationary and streaming data obtained from WBAN sensors; (6) it provides a transparent learning process. These features make VFDT a suitable candidate for implementing an autonomous decision maker for DDoS attack detection in cloud-assisted WBAN.

In this paper an improvement of VFDT [3], namely, enhanced VFDT (EVFDT), is proposed which differs from the existing algorithms in terms of classification accuracy, tree size, memory, and time. Our proposal is to build a decision tree based classification algorithm capable of handling noisy data and detecting a DDoS attack efficiently with high accuracy and low false alarm rate while allowing legitimate requesters to access the resources. The proposed algorithm is deployed at the victim node.

The contributions of this paper include the following:

(1) It proposed a novel EVFDT classification algorithm that is capable of classifying network traffic and detecting DDoS attack with high accuracy and less false alarms. EVFDT has achieved classification accuracy of about 96.5% with 0% noise and 81.5% with 20% noise in dataset.

(2) It is analyzing the effect of noise on stream data and its impact on the accuracy and false alarm rate while detecting a malicious behavior in the network.

(3) It highlights the shortcomings of the existing DDoS detection techniques.

(4) For simulation experiment, a DDoS attack algorithm is written and simulated in order to generate attack traffic for evaluation.

(5) It presents evaluation of the proposed algorithm and results.

The remainder of this paper is organized as follows: Section 2 discusses the existing data mining and stream mining techniques for detecting DDoS attack along with their limitation when applied in resource scarce WBAN environment. Section 3 elucidates the proposed system model and proposed algorithms. Section 4 presents the simulation experiment with proposed algorithm for generating DDoS attack strategy. At the end of Section 4, a detailed performance analysis and results comparison are given. Finally, Section 5 concludes the paper with recommendation for future work.

2. Related Work

2.1. Distributed Denial of Service (DDoS) Attack. A Distributed Denial of Service (DDoS) attack is defined as an explicit attempt by an attacker to exhaust the resources of a victim node. Multiple nodes are deployed to launch an attack by sending a stream of packets towards the victim, thus consuming the key resources of victim node and making them unavailable to legitimate nodes. These resources mainly include the network bandwidth, computing power, and memory resources [4].

DDoS attack is mainly divided into two classes [5], bandwidth depletion attack and resource depletion attack. In bandwidth depletion attack, the goal of an attacker is to flood the victim node with huge amount of traffic in order to prevent the legitimate traffic from reaching the victim node. It is further divided into flood attack and amplification attack. In resource depletion attack, the goal of an attacker is to degenerate the critical resources (processor and memory) of a victim node in order to prevent the legitimate user from using these resources.

A detailed analysis of DDoS attack in cloud-assisted WBAN environment and its implication shows that DDoS attack has following characteristics:

(1) During an attack, the packet length, sequence number, and window size remain fixed.

(2) Source IP and destination IP address along with port numbers are spoofed and generated randomly.

(3) Packet throughput decreases for legitimate users, which is defined as the number of bytes transferred from source to destination per unit time.

(4) Packet loss increases for legitimate users, which occurs due to the interaction of legitimate traffic with attack traffic.

(5) Packet delay increases as network congestion builds up.

(6) Traffic jitter may also increase significantly to impact especially real-time traffic.

Taking into account these characteristics, the key challenge lies in identifying the attack traffic from legitimate traffic of incoming data stream.

2.2. Data Mining Techniques. In the recent past, data mining techniques have been considered as one of the most promising solutions for identifying the malicious behavior of nodes in the network. For this research, data mining techniques have been studied and evaluated for the detection of DDoS attack in cloud-assisted WBAN environment. From the perspective of DDoS attack detection, existing data mining techniques (Subbulakshmi et al. [6], Wu et al. [7], Lee et al. [8], Arun and Selvakumar [9], and Thwe and Thandar [10]) can be broadly classified into source-based and destination-based detection techniques. Source-based detection techniques are deployed near the source of an attack whereas destination-based detection techniques are deployed near the victim of an attack.

Subbulakshmi et al. [6] proposed an Intrusion Detection System (IDS) based defense mechanism to counter DDoS attack. The IDS is trained using the datasets obtained from the extraction of attack traffic features. To strengthen the detection process, weights are added with these datasets at regular intervals.

Wu et al. [7] proposed a destination-based DDoS attack detection technique. In this technique, a decision tree is deployed for attack detection and a traffic pattern matching technique for attack identification and its traceback. For the classification of network traffic, fifteen different network and packet features are selected. C4.5 classification algorithm is deployed to classify the network traffic on the basis of identified packet features.

In [8], a source-based attack detection technique is proposed based on an enhanced traffic matrix approach. Traffic matrix parameters are optimized using a genetic algorithm (GA). To construct a traffic matrix, two features of the IP header are used, namely, the packet arrival time and source IP address. From the resultant traffic matrix, variance is calculated and used to classify the traffic as normal (high variance) or a DDoS attack (low variance). Finally alerts are generated upon the detection of an attack.

In [9], the author explores the ensemble based neurofuzzy classifier. The author contributes to existing classifier by including a weight update distribution policy, error cost reduction, and ensemble output combination approach. For performance evaluation, training and testing datasets are separately maintained. The result shows that this scheme efficiently detects DDoS attack.

Thwe and Thandar [10] proposed a statistical anomaly detection technique based on K-Nearest Neighbor (KNN) deployed at the victim node. A user specified threshold is defined. When the current state of the system differs from the defined model by a specified threshold, an anomaly is raised. At this stage KNN is used to detect an attack.

A detailed comparison of these techniques along with their limitations can be found in [11]. The major drawback of data mining techniques is that they are not suitable for real-time data mining of network traffic and require a huge amount of memory to store the datasets. In recent past, stream mining techniques have been proposed to overcome the limitations of data mining techniques.

2.3. Stream Mining Techniques: Very Fast Decision Tree (VFDT). Taking into account the resource constrained nature of WBAN sensors, VFDT proves to be a lightweight data mining technique that is able to process a large amount of high speed streaming data consuming less memory space. It turns out to be efficient in the detection of DDoS attack at any stage due to its ability of building decision tree from scratch. In [11] VFDT is applied for detection of DDoS attack and objective based comparison is done. The results show that the VFDT proves to be an accurate tool for DDoS attack detection. Therefore, VFDT is selected and improved for detecting DDoS attack efficiently.

2.3.1. Variants of VFDT. VFDT is a stream-based data classification method that learns using a complete set of N training samples expressed as (X, y), where X is a vector of n attributes given as $\{X_1, X_2, \ldots, X_n\}$. The aim is to construct a model of a mapping function $y = f(X)$ that will predict the classes of subsequent samples x with maximum accuracy. To design a VFDT for DDoS attack detection, the mathematical preliminaries used for the classification are first discussed (see [3, 12]).

Hoeffding Bound. This gives a certain level of confidence about the best attribute to split the node. Suppose we have N independent observations of a real-valued random variable r whose bounded range is R. The Hoeffding bound states that, with confidence level $1 - \delta$, the true mean of variable r is at least $r - \epsilon$, where ϵ can be calculated using

$$\epsilon = \sqrt{\frac{R^2 \ln(1/\delta)}{2N}}. \tag{1}$$

Information Gain. VFDT uses the information gain as heuristic evaluation function to find the upper and lower bounds with high confidence. The upper bound $G(\cdot)^+$ and lower bound $G(\cdot)^-$ are calculated using

$$G(A,T)^+ = \sum_{v \in A} P(T,A,v) + \sqrt{\frac{\ln(1/\delta)}{2N}} H(\mathrm{Sel}(T,A,v))^+$$
$$G(A,T)^- = \sum_{v \in A} P(T,A,v) + \sqrt{\frac{\ln(1/\delta)}{2N}} H(\mathrm{Sel}(T,A,v))^-, \tag{2}$$

where A is an attribute in the T set of training samples. $P(T, A, v)$ is a fragment of the training samples in set T that holds the value v for attribute A. $\mathrm{Sel}(T, A, v)$ selects all the training samples having value v for attribute A from set T.

Although VFDT classifies stream data efficiently, it has limitations like the fact that it cannot handle noisy data and classification accuracy decreases with the increase in noise. In recent past, few variations of VFDT have been proposed. In this section, variants of VFDT are discussed and briefly analyzed for their feasibility long with their limitations when employed for DDoS attack detection in cloud-assisted WBAN. The variations of VFDT are analyzed on the following parameters: attack detection accuracy, time, memory, and tree size.

Domingos and Hulten [3] proposed VFDT based on Hoeffding bound (HB) using (1) to control over error in the attribute splitting distribution selection. Information gain $G(\cdot)$ given in (2) is used as heuristic evaluation function (HEF) in order to decide the split attribute to convert the decision nodes to leaves. $\Delta G = G(X_a) - G(X_b)$ defines the difference between the two best attributes X_a (highest $G(\cdot)$) and X_b (second highest $G(\cdot)$). If $\Delta G > \epsilon$, then X_a is considered as highest value attribute in $G(\cdot)$. At this stage, the splitting occurs on attribute X_a and the decision node is converted into leaf node. The major drawback of this technique lies in the fact that there exist certain cases, when the two information gains have very small differences and are equally good to become a leaf node. At this stage, a tie condition occurs

and the process gets stuck. Resolving the tie-breaking is a computation intensive task that increases the processing time and decreases the overall accuracy of decision tree. At the same time, it is considered to be inappropriate for resource constrained WBAN.

To overcome the limitation of VFDT, Hulten et al. [12] proposed a fixed tie-breaking threshold τ. Whenever the difference between two information gains is very small, τ acts as a quick decisive parameter to solve the tie condition. The node splitting occurs on the current best attribute regardless of how good the second best attribute might be. The value of τ is chosen randomly and remains fixed throughout the process. An excessive tie-breaking condition reduces the performance of VFDT-τ significantly on noisy and complex streaming data, even with the use of parameter τ. VFDT-τ does not support pruning as the tree size itself is very small. While improving the accuracy, the tree size explodes. Therefore, a suitable pruning mechanism is required at this stage.

To overcome the limitation of fixed tie-breaking threshold, Yang and Fong [13] proposed an algorithm based on adaptive tie-breaking threshold computed directly from the Hoeffding bound mean. The value of HB mean fluctuates intensively with the increase in the noise percentage thus reducing the accuracy of attack classification.

Concept Adaptive VFDT (CVFDT) [12] maintains two trees simultaneously in memory. The tree with the shortest depth is retained and the other one is discarded. The main drawback of this technique is that it consumes more memory and time to maintain two trees. Also CVFDT does not handle noisy data efficiently.

2.4. Effect of Noise in Streaming Data. Noisy data is considered as meaningless or extraneous data that makes the identification of data patterns more difficult. As the noise increases in data stream, the number of outliers also increases. In sensor networks, noise arises due to the changes in system behavior and malicious activity in the network. There are two major sources of noise [14].

Error. An error is defined as a noisy value coming from an erroneous sensor. Outliers caused by such errors have a very high probability of occurrence.

Event. An event refers to a particular phenomenon which, in this case, is an attack occurrence event that changes the state of a system. Outliers caused by events occur with small probability but they are lasting and modify the historical patterns of sensor data.

In both cases, the presence of outliers due to noisy data decreases the attack detection accuracy and increases the false alarm rate. At the same time, the tree size increases which results in added memory consumption. The goal is to remove these outliers from the sensor data in order to ensure the high detection accuracy while keeping the resource consumption of the network minimal.

Figure 1 shows the detrimental effect of noisy data on classification accuracy (Figure 1(a)) and tree size (Figure 1(b)). The experimental data is synthetic and supplied as input to VFDT-τ algorithm with $\tau = 0.05$ [12]. The error rate significantly affects both the accuracy and tree size of decision tree when the number of instances increases manifold. Even a small error rate leads to increase in tree size by several times.

3. Proposed Model

3.1. System Architecture. The proposed Distributed Denial of Service (DDoS) attack detection system studies the network traffic behavior and classifies it as a normal or malicious traffic based on the observed traffic patterns. The proposed system architecture is shown in Figure 2. It consists of following phases starting from data collection phase up to response generation phase.

3.1.1. Data Collection Phase. In this phase, the incoming data stream is captured online and stored in database for training purposes. The captured data is supplied as an input to the preprocessing phase for feature extraction. Each instance of an incoming traffic is defined by a collection of features and is represented in feature vector space.

3.1.2. Preprocessing Phase. Preprocessing phase is further divided into the packet feature extraction phase followed by the labeling phase. In feature extraction phase, the real-time packets are captured from the network traffic in order to construct the new statistical features. These statistical features are listed in Table 1 and are used for DDoS attack detection and analysis. The identified features are important in defining the quality of service (QoS) of the network and to classify the network traffic pattern under DDoS attack. In labeling phase, classes are assigned to these statistical features. The entire dataset is divided into two classes labeled as "1" for attack and "0" for nonattack packet. After labeling the resulting dataset consists of both attack and nonattack data and is used for training the classification tree.

Mapping the preprocessing phase to feature vector space is given as follows:

(i) Let "x" be the N-dimensional vector of extracted features; that is, $x = \{x_1, x_2, x_3, \ldots, x_n\}$, where $\{1, 2, 3, \ldots, n\}$ are the individual packet features.

(ii) Let P_x be the packet of x features.

(iii) Let C_x be the vector space of labeled packets of dimension C_n.

3.1.3. Attack Classification. In this phase, the incoming traffic is classified as attack or nonattack by building a classification tree using the preprocessed data defined in feature vector space. For building the classification tree, we have proposed an algorithm, which is discussed in the next section.

3.1.4. Attack Response. The goal of attack response module is to minimize the impact of DDoS attack on the victim node while allowing the legitimate traffic to move forward. When a DDoS attack is detected, an appropriate traceback mechanism is applied to trace an attacker and block his traffic. The traceback technique will be explored in future work.

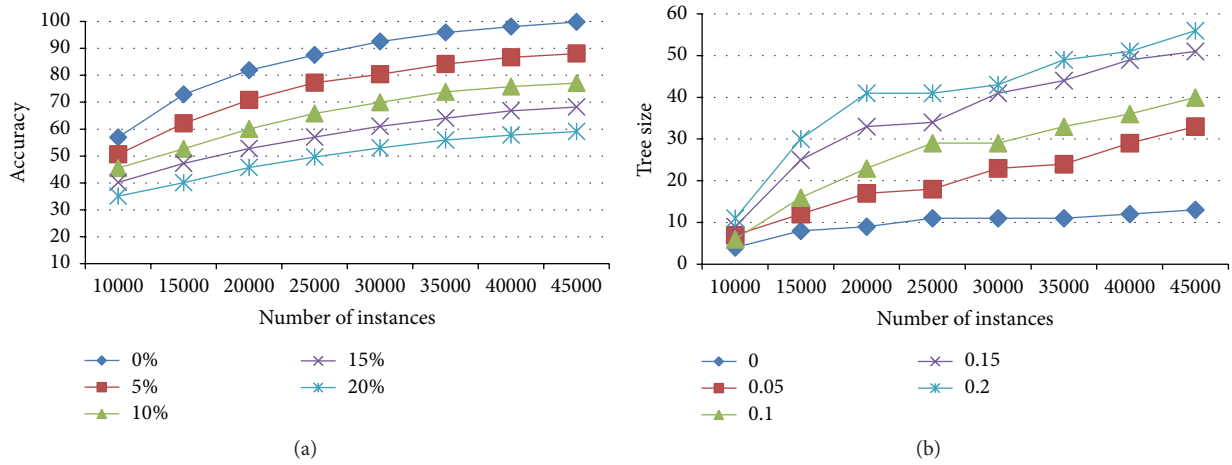


Figure 1: (a) Effect of noise on accuracy. (b) Effect of noise on tree size.

Table 1: List of features.

Features	Description	Formula for calculation
Packet loss (PL) percentage	Number of packets lost due to the interaction of the legitimate traffic with the attack. It is the presence of congestion in the network due to DDoS flooding attacks.	$\text{Loss} = \left[\frac{\sum \text{PL}}{\sum \text{PS}}\right] * 100.$ (i) PL is the packet lost. (ii) PS is the total number of packets sent towards the destination.
Delay or latency	The time taken by the packets to reach from source to destination.	$\text{Delay} = \sum (\text{PATime}_i - \text{PSTime}_i).$ (i) PATime_i is the packet arrival time. (ii) PSTime_i the packet start time.
Jitter (delay variation)	The variation in the time between packets arriving within a particular window. Jitter is used as an indicator of consistency and stability of network.	$\text{Jitter} = \sum_{i=0}^{n} \left[\frac{(\text{Delay}_i - \text{Delay})}{N}\right].$ (i) Delay_i is packet duration. (ii) Delay is the last packet delay. (iii) N is the difference of packet sequence number
Throughput (TH)	Bytes transferred per unit time from source to destination. The DDoS defense mechanism ideally increases the throughput for the legitimate users.	$\text{TH} = \frac{\sum_{i=0}^{n} \text{PacketReceived}}{\sum_{i=0}^{n} (\text{PacketStartTime} - \text{StopTime})}.$

3.2. Enhanced Very Fast Decision Tree (EVFDT): A Proposed Classification Algorithm. The proposed classification algorithm is an enhancement of original VFDT-τ [12] to make it efficient for the detection of DDoS attack in resource constrained cloud-assisted WBAN. The EVFDT classification algorithm simultaneously trains and tests the decision tree based on learning traffic patterns and classifies malicious behavior based on these learned patterns as shown in Figure 2.

Considering the severity of DDoS attack, the scarce resources, and missing protection mechanism of WBAN, the EVFDT improves the existing VFDT algorithms in terms of following parameters: (1) accuracy; (2) tree size; (3) time; (4) memory.

3.2.1. EVFDT Tree Building Process. EVFDT is based on original VFDT-τ [12] and is improved in two aspects: accuracy and the tree size. Algorithm 1 presents the pseudocode of EVFDT. It is divided into four subprocedures as described below.

(A) Procedure for Tree Initialization. In this procedure, the tree is initialized using a single leaf node, which is a root node. The tree grows as a new data stream arrives at the root node. Algorithm 2 shows the pseudocode of Tree Initialization Procedure. The procedure is executed when the first data stream arrives and it is the same as VFDT-τ [12].

(B) Procedure for New Stream Sample. In this procedure, data from the stream is traversed starting from the root node.

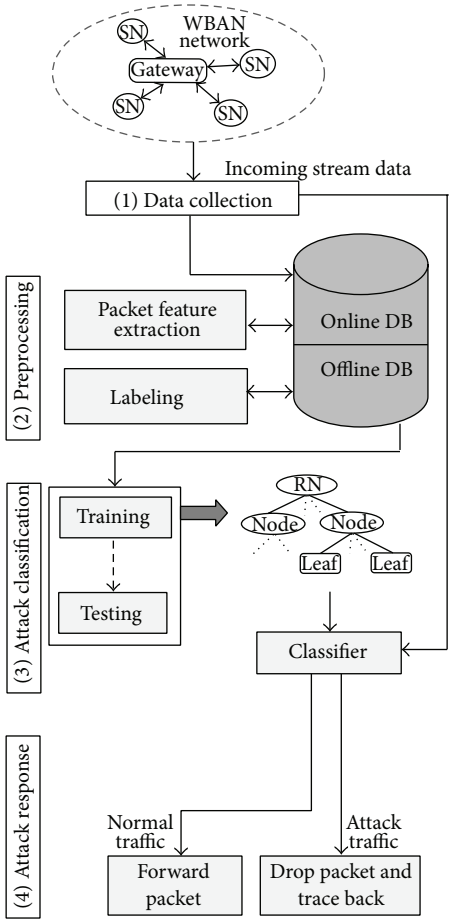


Figure 2: Proposed system architecture.

Each time a new data stream arrives, this procedure traverses the stream from root node to the leaf node and at the same time the tree statistics are updated. Algorithm 3 shows the pseudocode for traversing a new sample. This procedure is same as tree traversing of VFDT-τ [12].

(C) Procedure for Accuracy Improvement. Algorithm 4 shows the proposed pseudocode for EVFDT, that is, enhancing the accuracy of VFDT-τ [12]. VFDT-τ takes a fixed value of τ to break the tie as discussed in Section 2.3, which can be deviating because of presence of noise. To overcome this, optimized VFDT (OVFDT) takes the mean of Hoeffding bound (HB) values for tie-breaking. As noise ratio is high in sensor data, a simple mean does not depict the true mean of HB because of the presence of outliers. To cater for such outliers, EVFDT incorporates a procedure for accuracy enhancement as given in Algorithm 4 and explained as follows.

Let XS be the sorted list of HB values, m the total number of HB values in XS, and n the new HB value seen at the node. Starting with $m \geq 3$, the mean of difference between HB values in XS is calculated. This mean is stored as threshold T_r. Let XS_n be the position of n in XS and the PrunedMean is calculated with new value n. If this PrunedMean is less than the threshold T_r, then n is added to XS; otherwise, n is discarded as an outlier.

(D) Procedure for Tree Pruning. Algorithm 5 shows the pseudocode for the proposed tree pruning and explained as follows.

Let HT be the Hoeffding tree to be pruned and S the stream of examples belonging to S_0 and DataSeenAtLeafn_0 is the number of samples seen at leaf node. For every example S in S_0 is passed through the tree starting from the root node. If S_0 is filtered to the leaf node n_0, then increment DataSeenAtLeafn_0 otherwise continues growing the HT.

If DataSeenAtLeafn_0 is less than τ, where τ is the threshold for checking the eligibility of a node to be part of HT, then prune the tree by deleting the leaf node n_0 and updating the HT. The eligibility of the node to be part of HT is checked by comparing the number of samples seen at the leaf node with τ. The comparison will tell us that this leaf node has less contribution towards classification as a smaller number of instances are filtered to this leaf node on the current HT.

4. Simulation Experiments and Evaluation

In this section, we evaluate the performance of proposed EVFDT in detecting DDoS attack from incoming data stream in cloud-assisted WBAN environment. EVFDT is compared with VFDT and its variants in terms of the following parameters: accuracy, time, tree size, and memory. These evaluation parameters are important for comparison. Analysis shows that a higher classification accuracy produces a bigger tree size which in turn consumes more memory space and takes more learning time. Similarly, a faster learning speed brings a smaller tree size, which results in less memory space but lower classification accuracy.

The proposed EVFDT algorithm has been implemented using Very Fast Machine Learning (VFML) libraries [15]. The experiments are run on windows7-64-bit workstation with 2.8 GHz CPU and 8 GB RAM with all background processes switched off.

4.1. Synthetic Datasets. The Low Energy Adaptive Clustering Hierarchy (LEACH) protocol [16] was implemented in NS-2 for generating the synthetic data stream containing 1 million data values, which are divided into five datasets. Each dataset contains different number of instances and noise percentage. LEACH protocol is selected because it is lightweight and closely reflects the WBAN scenario. Cluster heads act as body control unit (BCU) and all other nodes as sensor nodes as depicted in Figure 3. LEACH is responsible for transferring the data from WBAN sensor node to BCU. The simulation runs for 900 s, 1200 s, 1500 s, 1800 s, and 2000 s to generate the data streams. Other simulation parameters and network configurations are shown in Table 2.

The resulting dataset includes both attack and nonattack data. The DDoS attack code is attached to sensor nodes to make them malicious. The number of malicious sensor nodes varies with different attack dataset.

Input:
S: a stream of examples
X: a set of symbolic attributes
$G(\cdot)$: heuristic evaluation function for node splitting
δ: one minus desired probability of choosing the correct attribute at any given node.
$n_{\min}$: number of samples between estimation of growth
XS: sorted list of Hoeffding bound values
m: total number of values in XS
n: new Hoeffding Bound value seen at the node
T_r: adaptive threshold
S_0: subset of $S \rightarrow S_0 \in S$
τ: 5% of examples in S_0. Threshold for checking the eligibility of a node to be part of HT
Size of $S_0 = n_{\min}$
Output:
A decision tree HT
Procedure EnhancedVFDT(S, X, G, δ, $n_{\min}$)
BEGIN:
A stream of examples S arrives
IF (HT = ϕ), THEN *TreeInitialization*(S, X)
Get an Initialized HT with a single root node
IF (HT ≠ ϕ), THEN *NewStreamSample*(S, X)
Label l with the majority class among the samples seen so far at l
Let n_1 be the number of samples seen at l
IF the samples seen so far at l are not all of the same class and (n_l mode $n_{\min}$) = 0
THEN
Compute $G_l(X_i)$ for each attribute $X_i \in X_l - \{X_\phi\}$ using $n_{ijk}(l)$
PrunedMean = *AccuracyEVFDT*(XS, m, n, T_r)
Let X_a be the attribute with highest G_l and X_b be the attribute with second-highest G_l
Compute ε using (1)
Let $\Delta G_l = G_l(X_a) - G_l(X_b)$
IF ((ΔG_l) > ε or (ΔG_l) ≤ *PrunedMean*) and $X_a \neq X_\phi$, THEN split X_a as a branch
FOR each branch of split
Add a new leaf l_m and let $X_m = X - \{X_a\}$
Let $G_m(X_\phi)$ be the G obtained by predicting the most frequent class at l_m
FOR each class y_k and each value x_{ij} of each attribute $X_i \in X_l - \{X_\phi\}$
Let $n_{ijk}(l_m) = 0$.
END-FOR
END-FOR
END-IF
ELSE *Pruning*(S, S_0, $n_{\min}$, τ, HT)
Return HT
END:

Algorithm 1: EVFDT procedure: enhanced EVFDT.

Procedure TreeInitialization(S, X)
BEGIN:
Let HT be a tree with a single leaf l_1 (the root)
Let $X_1 = X \cup (X_\phi)$
Let $G_1(X_\phi)$ be the G obtained by predicting the most frequent class in S
FOR each class y_k
FOR each value x_{ij} of each attribute $X_i \in X$
Let $n_{ijk}(l) = 0$
END-FOR
END- FOR
Return HT
END:

Algorithm 2: EVFDT procedure: tree initialization.

Procedure NewStreamSample(S, X)
BEGIN:
  FOR each new instance (x, y) in S
    Sort (x, y) into a leaf l using HT
  FOR each x_{ij} in X such that $X_i \in X$
    Increment $n_{ijk}(l)$
  END-FOR
END:

Algorithm 3: EVFDT procedure: new stream sample.

Table 2: Simulation parameters.

Parameters	Values
Sensing field	50 × 50
Topology	Star
Simulation time	900 s, 1200 s, 1500 s, 1800 s, 2000 s
Packet size	1000 bytes
Radio communication range	2 m standard, 5 m special use
Number of nodes	50
BAN coordinator	Directional mode
Sensor nodes	Omnidirectional mode
Routing protocol	LEACH
Transport protocol	TCP

4.1.1. DDoS Attack Strategy: Generation and Analysis. In this section, an attack dataset is generated and analyzed at TCP layer. In ongoing simulation experiment, first the normal TCP traffic was considered for analysis. Secondly, the attack was generated and its intensity under flooding attack with TCP traffic was analyzed. An attack algorithm was written and generated online. The resulting dataset (attack dataset) is stored in database at base station for preprocessing. The proposed EVFDT algorithm was applied on the preprocessed dataset. The DDoS attack algorithm is presented in Algorithm 6.

Java code is written to randomly add noise in the dataset. For this purpose, N-dimensional feature vector is multiplied with vector of random variables taken from the Normal Distribution $N(0, \sigma 2)$, where $\sigma 2$ is noise variance adjusted according to the percentage of noise added.

To evaluate the performance of EVFDT, noise is attached to datasets in order to compare the result of EVFDT with VFDT and its variants under different noise percentage. Each dataset contains different number of instances and divided into 20% testing data and 80% training data to learn the classifier. The performance is evaluated using the following parameters.

(1) Attack Detection Accuracy. In general, the accuracy of EVFDT is directly proportional to the number of data stream samples. With the increase in data stream samples, the EVFDT becomes more and more accurate as long as there is noise-free data. But as soon as the noise is injected, the classification accuracy starts decreasing as shown in Figure 4.

For the conducted simulation experiments, attack detection accuracy was calculated using

$$\text{Detection Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total number of Tested Examples}}, \quad (3)$$

where True Positive (TP) is number of samples correctly classified as attack class, True Negative (TN) is number of samples correctly classified as nonattack class, False Positive (FP) is number of samples incorrectly classified as attack class, and False Negative (FN) is number of samples incorrectly classified as nonattack class.

Figure 5 shows the accuracy comparison of EVFDT with existing VFDT and its variants on datasets with noise percentage of 10% and 20%. On all experimental datasets, EVFDT maintains higher accuracy with less false alarm rate and more sensitivity.

Table 3 shows the sensitivity, specificity, and false alarm rate (FAR) of algorithms with respect to different noise percentages. These can be calculated using (4) (sensitivity), (5) (specificity), and (6) (FAR)

$$\text{Sensitivity} = \frac{(\text{True Positives})}{(\text{True Positives} + \text{False Negatives})}, \quad (4)$$

$$\text{Specificity} = \frac{(\text{True Negatives})}{(\text{True Negatives} + \text{False Positives})}, \quad (5)$$

$$\text{FAR} = \frac{(\text{False Positives})}{(\text{False Positives} + \text{True Negatives})}. \quad (6)$$

Results show that the average accuracy of EVFDT is greater than VFDT and its variants. VFDT-τ has lowest accuracy among all algorithms.

(2) Tree Size. A significant characteristic of EVFDT lies in its ability to build a decision tree with reduced tree size and increased classification accuracy simultaneously. The tree size gets bigger with increase in noise percentage as shown in Figure 6. Although VFDT-τ obtains smallest tree size in our simulation, it results in increased classification error. As shown in Table 4, notably, EVFDT maintains a smaller tree size. In few cases, EVFDT and VFDT-τ produce same tree size, but the average tree size of EVFDT is smaller than VFDT-τ. The tree size is the depth of the decision tree. The simulation experiment shows the tree size (TS): $\text{TS}_{\text{EVFDT}} < \text{TS}_{\text{VFDT-}\tau} < \text{TS}_{\text{CVFDT}} < \text{TS}_{\text{OVFDT}}$.

(3) Computation Time. Early detection of an attack is a desirable property of any algorithm. A detection mechanism should be fast enough in detecting an attack. Computational time is the total time taken in seconds for processing a full set of data stream. Computational complexity of proposed algorithm is proportional to $O(lpdvc)$, where lp is the length of a pruned tree, d is the total number of attributes, v is the number of values per attribute, and c is the number of classes. VFDT-τ has a small running time due to small and fixed value

Procedure AccuracyEVFDT(XS, m, n, T_r)
BEGIN
IF ($m == 1$)
$T_r = \sum \frac{XS}{m}$
ELSE-IF ($m == 2$)
$T_r = \sum \frac{XS}{m}$
FOR ($j = 1$)
$XD_j = XS_{j+1} - XS_j$
Increment $j = j + 1$
$j < m$
END-FOR
ELSE // **Calculate mean difference of** XS **without** n
FOR ($j = 1$)
$XD_j = XS_{j+1} - XS_j$
Increment $j = j + 1$
$j < m$
END-FOR
$T_r = (\sum_{i=1}^{m} XD_i)/m$ // **Updated Threshold**
Let XS_n be the position of "n" in sorted list
$$\text{PrunedMean} = \frac{(XS_n - XS_{n-1}) + (XS_{n+1} - XS_n)}{2}$$
IF (PrunedMean $\leq T_r$)
THEN
Add n in sorted list XS
$XS \longleftarrow XS + n$
ELSE
Discard n as it is outlier
$XS \longleftarrow XS$
Return **PrunedMean**
END

Algorithm 4: EVFDT procedure: accuracy EVFDT.

Procedure Pruning($S, S_0, n_{\min}, \tau$, HT)
BEGIN:
Let DataSeenAtLeaf be the number of samples seen at leaf node n_0
FOR each example $S \in S_0$
IF the sample S traverses to the node n_0 which is a leaf node
THEN
Start counter on node n_0
Increment: DataSeenAtLeaf n_0
ELSE
Continue growing EVFDT
END-FOR
IF (DataSeenAtLeaf $< \tau$)
THEN
Prune the tree: Delete n_0
UPDATE HT
END:

Algorithm 5: EVFDT procedure: pruned EVFDT.

τ. EVFDT takes slightly more time than VFDT-τ because of pruning and threshold computation as shown in Figure 7. The computation time (CT) is compared as $\text{CT}_{\text{VFDT-}\tau} < \text{CT}_{\text{EVFDT}} < \text{CT}_{\text{OVFDT}} < \text{CT}_{\text{CVFDT}}$.

(4) Memory. Taking into account the resource scarcity of WBAN environment, the proposed mechanism should consume less memory resources as shown in Figure 8. The advantage of VFDT over existing machine learning techniques is

Input:
SN: Set of sensor nodes for cluster head selection
R: Number of rounds
Simulation Parameters of LEACH protocol
Output:
DDoS Attack Dataset
Procedure DDoSAttackAlgorithm(SN, R)
BEGIN:
IF ($r = 0$)
Initial round CH selection
FOR maximum number of rounds "r"
Choose r rounds CH randomly
CH announces schedule time T to all SN
END-FOR
Attach attack code with random nodes N_r
Randomly initiates malicious nodes towards victim node V
Malicious nodes start and stop randomly according to time T during formation
Malicious nodes start compromise victim node V
Malicious nodes forward flooding packets to V with high rate to overflow and consume resources available to victim
Victim node V receives packet with high rate
More malicious nodes start compromising victim at their schedule time T causing DDoS flooding attack
END:

Algorithm 6: DDoS attack algorithm.

Table 3: Sensitivity, specificity, and FAR of the algorithms.

Noise	VFDT-τ (%)				CVFDT (%)				OVFDT (%)				EVFDT (%)			
	SN	SP	AC	FA	SN	SP	AC	FA	SN	SP	AC	FA	SN	SP	AC	FA
0%	90.8	98.4	92.2	5.4	89.8	99.4	94.6	5.1	97.6	94.9	96.2	2.4	97.9	94.7	96.5	1.1
5%	79.2	91.7	87.4	5.9	78.9	91.7	87.5	5.7	79.2	91.7	87.6	2.7	82.3	90.0	89.8	1.3
10%	78.2	88.8	82.3	6.8	69.2	90.8	79.6	6.2	77.0	90.7	83.6	3.2	80.2	88.4	84.2	1.7
15%	76.3	88.3	81.2	7.3	67.7	88.9	78.0	7.0	76.9	86.8	80.4	3.5	79.2	86.8	83.0	2.2
20%	77.7	81.3	78.2	7.9	66.9	79.1	76.9	7.8	79.9	81.0	79.3	4.1	78.1	83.4	81.5	2.7
Avg.	**80.4**	**88.9**	**84.2**	**6.7**	**74.5**	**89.1**	**83.3**	**6.3**	**82.5**	**89.0**	**85.4**	**3.2**	**83.6**	**88.6**	**87.0**	**1.8**

Table 4: Tree size comparison with different noise percentage.

Dataset	Noise	VFDT-τ	CVFDT	OVFDT	EVFDT
Dataset-10	0%	4	5	6	3
	5%	6	7	8	6
	10%	9	9	11	8
	15%	9	10	13	9
	20%	9	9	15	8
	Average	**7**	**8**	**9**	**6**
Dataset-10	0%	5	6	5	5
	5%	7	7	8	7
	10%	8	9	10	6
	15%	10	11	11	8
	20%	10	11	13	8
	Average	**8**	**9**	**10**	**6**

that it does not require a full dataset to be stored in memory. Memory required for running EVFDT is proportional to $O(ndvc)$, where n is the number of decision nodes in a tree, d is the total number of attributes, v is the number of values per attribute, and c is the number of classes. The total amount of memory required to run EVFDT is the sum of memory allocated for learning and memory allocated for training. EVFDT consumes less memory than existing algorithms and is compared as (M): $M_{\text{EVFDT}} < M_{\text{OVFDT}} < M_{\text{CVFDT}} < M_{\text{VFDT-}\tau}$.

4.2. Comparison of EVFDT with VFDT and Its Variants. Comparison of EVFDT classification algorithm with existing VFDT and its variants is shown in Table 5. Only OVFDT and EVFDT can handle noisy data. At the same time, OVFDT handles noisy data to some extent and becomes inaccurate with the increase in noise percentage due to the presence of outliers. VFDT-τ, CVFDT, and IVFDT do not provide tree pruning. They maintain small tree size from the beginning but with increased error percentage in classification. Only proposed EVFDT algorithm efficiently handles noisy data and at the same time maintains small tree size with less resource usage.

Table 5: Comparison of existing machine learning techniques with VFDT.

Features	VFDT-τ	CVFDT	OVFDT	EVFDT
Detection accuracy	Very low	Very low	Good; does not handle outliers	Excellent
Resource usage (time/memory)	Less time; more memory	More time in building two trees; requires additional memory	Less time; less memory	Having same time as VFDT but consuming very less memory space
Noisy data handling	Does not handle noisy data	Not appropriate under noisy data	HB fluctuation intensifies under noisy data; accuracy decreases	Handles noisy data efficiently
Tree size/pruning	Small tree size; no pruning	Same tree size as VFDT; no pruning	Small tree size; incremental pruning	Small tree size; iterative pruning
Computational resources	Consuming less resources	Consuming more resources by maintaining two trees	Consuming less resources	Consuming very less resources by cutting of HB outliers

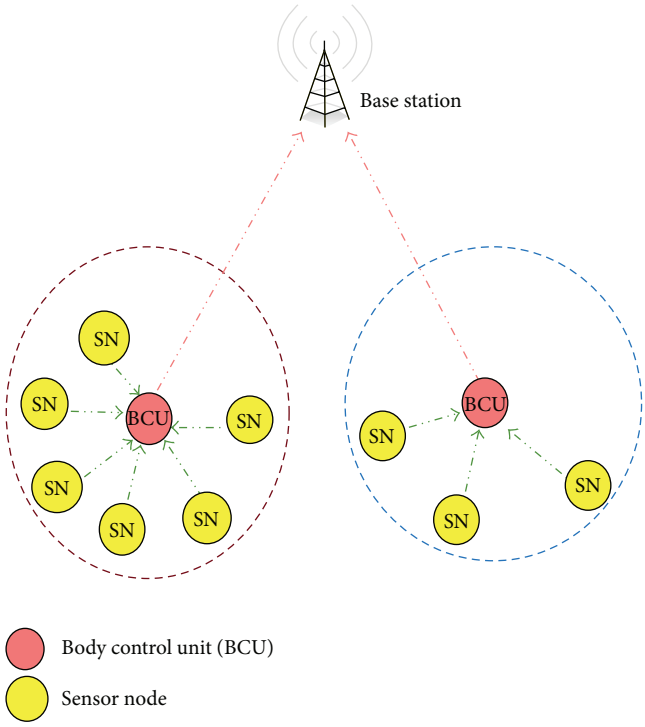


Figure 3: Illustration of LEACH protocol.

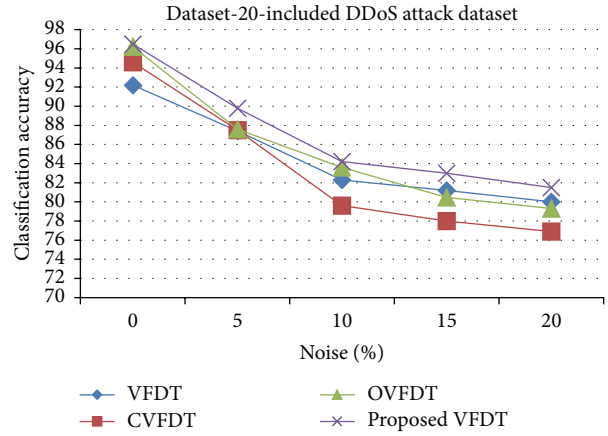


Figure 4: Accuracy in different noise percentage.

5. Conclusion and Future Work

Nowadays, zombies-based DDoS attack occurs with legitimate flow of traffic. Therefore, it is very difficult to detect such attacks even with the presence of stored attack traffic signatures. The challenge is to distinguish the legitimate traffic and DDoS attack traffic. Data mining techniques for data classification fail for real-time streaming data and also they require a sufficient amount of memory for data storage. On the other hand, stream mining techniques handle high speed streaming data originating from WBAN sensors and are efficient for resource scarce WBAN network. An algorithm based on VFDT is proposed in this paper. Our main contributions include a novel enhanced Very Fast Decision Tree (EVFDT) classification algorithm and it differs from existing algorithms in terms of attack classification accuracy and tree size. The performance of EVFDT algorithms is evaluated on the synthetic dataset generated by implementing a LEACH protocol. An attack code is written to generate DDoS attack. Experimental result shows that the proposed algorithm is able to detect an attack with high classification accuracy (96.5%) and low false alarm rate (1.1%) with less memory overhead. The proposed model is deployed at victim node and is deployed in real-time network environment. In future, the proposed architecture is deployed on real WBAN test bed to evaluate the performance of proposed algorithm.

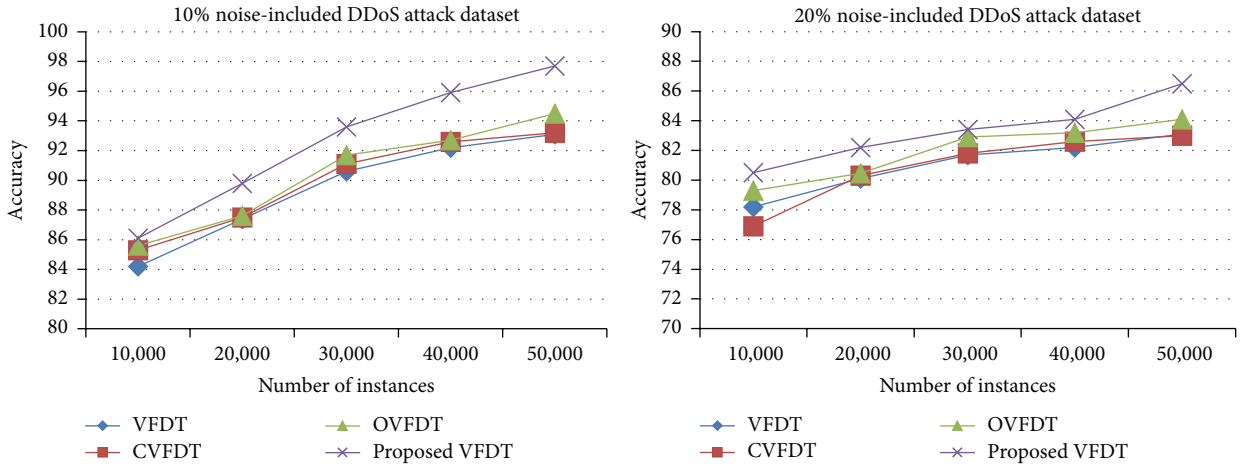


Figure 5: Accuracy versus number of instances in different noise ercentages.

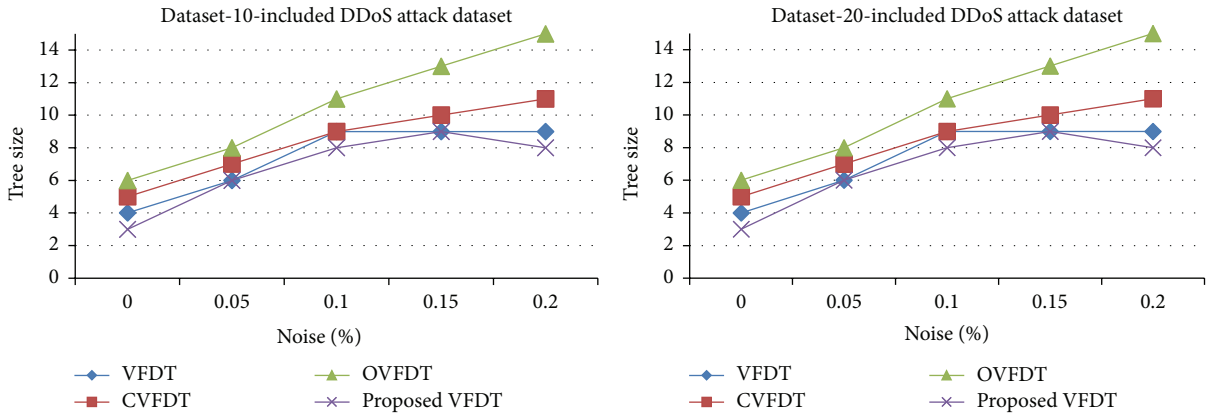


Figure 6: Tree size comparison with different noise percentage.

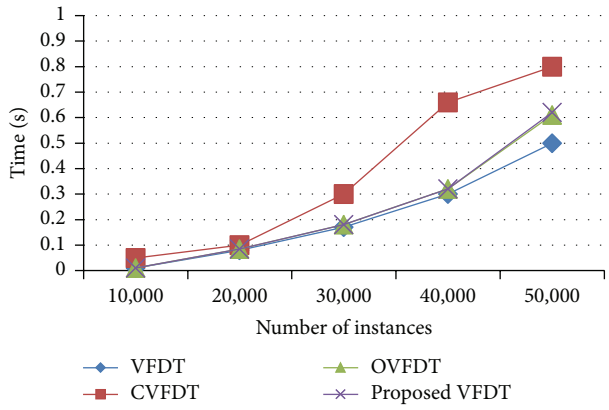


Figure 7: Computation time comparison.

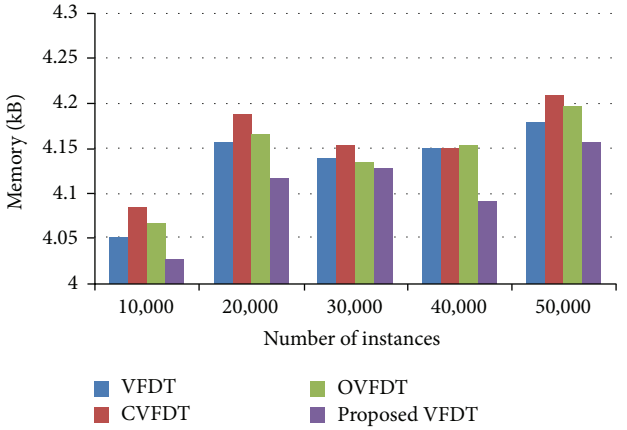


Figure 8: Memory resource comparison.

Conflict of Interests

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

Acknowledgments

The authors would like to extend their sincere appreciation to the Deanship of Scientific Research at King Saud University

for its funding of this research through the Research Group Project no. RG-1435-048. The authors would also like to thank the National University of Sciences and Technology, Islamabad, Pakistan, for its support during this research.

References

[1] R. Latif, H. Abbas, S. Assar, and S. Latif, "Analyzing feasibility for deploying very fast decision tree For DDoS attack detection in cloud-assisted WBAN," in *Intelligent Computing Theory: Proceedings of the 10th International Conference, ICIC 2014, Taiyuan, China, August 3-6, 2014*, Lecture Notes in Computer Science, pp. 507–519, Springer, Berlin, Germany, 2014.

[2] R. Latif, H. Abbas, and S. Assar, "Distributed denial of service (DDoS) attack in cloud- assisted wireless body area networks: a systematic literature review," *Journal of Medical Systems*, vol. 38, article 128, 2014.

[3] P. Domingos and G. Hulten, "Mining high-speed data streams," in *Proceedings of the 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 71–80, Boston, Mass, USA, August 2000.

[4] S. T. Zargar, J. Joshi, and D. Tipper, "A survey of defense mechanisms against distributed denial of service (DDOS) flooding attacks," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 4, pp. 2046–2069, 2013.

[5] D. Arora, P. Singh, and V. Singh, "Impact analysis of denial of service (DoS) due to packet flooding," *International Journal of Engineering Research and Applications*, vol. 4, no. 6, pp. 144–149, 2014.

[6] T. Subbulakshmi, S. M. Shalinie, V. Ganapathisubramanian, K. Balakrishnan, D. Anandkumar, and K. Kannathal, "Detection of DDoS attacks using enhanced support vector machines with real time generated dataset," in *Proceedings of the 3rd International Conference on Advanced Computing (ICoAC '11)*, pp. 17–22, IEEE, Chennai, India, December 2011.

[7] Y.-C. Wu, H.-R. Tseng, W. Yang, and R.-H. Jan, "Ddos detection and traceback with decision tree and grey relational analysis," *International Journal of Ad Hoc and Ubiquitous Computing*, vol. 7, no. 2, pp. 121–136, 2011.

[8] S. M. Lee, D. S. Kim, J. H. Lee, and J. S. Park, "Detection of DDoS attacks using optimized traffic matrix," *Computers and Mathematics with Applications*, vol. 63, no. 2, pp. 501–510, 2012.

[9] R. K. Arun and S. Selvakumar, "Detection of distributed denial of service attacks using an ensemble of adaptive and hybrid neuro-fuzzy systems," *Computer Communications*, vol. 36, no. 3, pp. 303–319, 2013.

[10] T. Thwe and P. Thandar, "Statistical anomaly detection of DDoS attacks using K-nearest neighbour," *International Journal of Computer & Communication Engineering Research*, vol. 2, no. 1, pp. 315–319, 2014.

[11] R. Latif, H. Abbas, and S. Assar, "Distributed Denial of Service (DDoS) attack in cloud—assisted wireless body area networks: a systematic literature review," *Journal of Medical Systems*, vol. 38, article 12, 2014.

[12] G. Hulten, L. Spencer, and P. Domingos, "Mining time-changing data streams," in *Proceedings of the 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '01)*, pp. 97–106, San Francisco, Calif, USA, August 2001.

[13] H. Yang and S. Fong, "Moderated VFDT in stream mining using adaptive tie threshold and incremental pruning," in *Proceedings of the 13th International Conference on Data Warehousing and Knowledge Discovery (DaWaK '11)*, pp. 471–483, Toulouse, France, August 2011.

[14] A. Fawzy, H. M. O. Mokhtar, and O. Hegazy, "Outliers detection and classification in wireless sensor networks," *Egyptian Informatics Journal*, vol. 14, no. 2, pp. 157–164, 2013.

[15] VFML (Very Fast Machine Learning) toolkit, 2014, http://www.cs.washington.edu/dm/vfml/.

[16] L. Hughes, X. Wang, and T. Chen, "A review of protocol implementations and energy efficient cross-layer design for wireless body area networks," *Sensors*, vol. 12, no. 11, pp. 14730–14773, 2012.