Republic of Iraq
Ministry of Higher Education and Scientific Research
University of Babylon
College of Sciences for Women
Department of Computer Sciences

Efficient EEG Data Compression Approaches for Fog Computing Based IoT Applications

A Thesis

Submitted to the Council of College of Science for Women, the University of Babylon in a Partial Fulfilment of the Requirements for the Degree of Master in Science\ Computer Sciences

By

Marwa saeed khlaif

Supervised By

Prof. Dr. Ali Kadhum Idrees

2022 A. D. **1444 A. H.**

بسم الله الرحمن الرحيم

ٱقۡرَأۡ بِٱسۡمِ رَبِّكَ ٱلَّذِی خَلَقَ ١

خَلَقَ ٱلۡإِنسَـٰنَ مِنۡ عَلَقٍ

٢ ٱقۡرَأۡ وَرَبُّكَ ٱلۡأَكۡرَمُ ٣

ٱلَّذِی عَلَّمَ بِٱلۡقَلَمِ ٤

عَلَّمَ ٱلۡإِنسَـٰنَ مَا لَمۡ يَعۡلَمۡ ٥

صَدقَ اللهُ العَظِيمْ

سُورَةُ العَلَقِ

الآيات (1-5)

Supervisor Certification

I certify that this thesis entitled

" Efficient EEG Data Compression Approaches for Fog Computing Based IoT Applications"

written by

"Marwa saeed khlaif"

was prepared under my supervision at College of Sciences for Women as a partial fulfillment of the requirements for the degree of a Master's in Computer Sciences.

Signature:

Name: Prof. Dr. Ali Kadhum Idrees

Date: / / 2022

Address: University of Babylon/College of Sciences for Women

Head of the Department Certification

In view of the available recommendations, I forward the thesis entitled "Efficient EEG Data Compression Approaches for Fog Computing Based IoT Applications" for debate by the examining committee.

Signature:
Name: Asst. Prof. Dr.saif Mahmoud Khalaf
Date: / / 2022
Address: University of Babylon/College of Science for Women

Acknowledgements

First of all, I thank God who inspired me with patience and strength to complete this study.

It is not easy except for what God makes easy.

I would like to express my sincere thanks and appreciation to the Supervisor

"Prof. Dr. Ali Kadhum Idrees"

To guide and follow up with me and provide important tips and suggestions to improve this study. Without him, I would not have completed this thesis. I learned a lot from him on this journey of research. I am really grateful to him.

Dedications

To my great creator, dear God

To the first teacher of mankind, the prophet Mohammed , may God bless him and his family and grant them peace.

To my intercession with God in this world and the hereafter, the pure imams, peace be upon them.

To the example of dedication and devotion .... my beloved father.

To whom offered me happiness and comfort over her happiness ... My honorable mother.

To those who supported me and encouraged me with all love and patience.........my husband and daughters

To those who wish happiness and success for me from the bottom of their hearts without any compensation ... my dear brothers and sisters.

To all my dear loyal friends

I offer you that humble work ....

Marwa

Abstract

The Internet of Medical Things (IoMT), which includes many innovative applications that connect to healthcare systems through computer networks over the Internet, has emerged due to the rapid growth of medical equipment and communication technology. These applications will regularly generate huge amounts of data, which the edge gateway will transfer to the cloud on a continuous and regular basis for further processing. The Internet of Things will be very taxing if this huge amount of data is sent through it to the cloud. The amount of data transmitted and prolonged processing delays greatly affect the reaction speed of IoT applications. This will reduce the response times of these applications. So to reduce the amount of data transferred and improve reaction times, IoT applications will benefit from the advantages of fog computing, which acts as an intermediate level between smart devices and the Internet.

The current thesis proposes lossless EEG data compression for IoT fog computing applications. Three encoding approaches are provided that can be performed at the fog gateway to reduce the patient EEG data sent to the cloud data center. These approaches are: First, lossless EEG data compression (LEDaC) technology from In order to reduce the amount of IoMT data, LEDaC combines two effective techniques of DBSCAN Clustering and Huffman encoding. Secondly, an Efficient Compression Technique (ECoT) is proposed for lossless EEG data compression. In order to communicate only the necessary EEG data to the cloud, ECoT approach integrates three effective methods into the fog gateway: DBSCAN Clustering, Delta encoding and Huffman encoding. Thirdly, A novel Lossless EEG Data Compression (NoLEDaC) approach is proposed to compress EEG data in fog computing for Internet of Medical things networks. This technique

reduces the size of the patient's EEG data by compressing it in a fog node before uploading it to the cloud data center.The proposed NoLEDaC approach combines: DBSCAN Clustering followed by RLE compression technique and Huffman encoding at the fog gateway.

Several experiments were completed successfully using (Bonn) dataset. The results showed that the proposed lossless compression approaches are ortperform other techniques in terms of compression ratio, compression time, decompression time, size of transmitted data, compression power, and average compression power. The proposed LEDaC approach offers a good compression ratio, compressing EEG data for each period from 2.1 up to 4.39 for different data set records, while ECoT approach compresses EEG data from 65% to 85% for different records. While EEG data was compressed by NoLEDaC approaches and obtained higher compression ratio than ECoT and LEDaC with an average of 66% to 98.6% in all records.

Table of Contents

Table of Figures

List of Tables

List of Abbreviations

Abbreviations	Meaning
ECoT	Efficient Compression Technique
EEG	Electroencephalogram
ECG	Electrocardiography
EMG	Electromyography
IoT	Internet of Things
IoMT	Internet of Medical Thinge
HE	Huffman Encoding
HCHE	Hierarchical Clustering and Huffman Encoding
HCLZW	Hierarchical Clustering and Lempel-Ziv-Welch
LEDaC	Lossless EEG Data Compression
LZW	Lempel-Ziv-Welch
NoLEDaC	Novel Lossless EEG Data Compression
WBSNs	Wireless Body Sensor Networks

Chapter One

General Introduction

1.1 Introduction

Wireless IoT (Internet of Things) devices like smart objects, sensors, cameras, and wearable devices are used in a variety of real-world applications, including smart homes, security, agriculture, smart cities, smart health care, military, and smart transportation [1]. The Internet of Things (IoT) is a standard that connects smart things to the Internet, allowing them to acquire, compute, and share data without the need for human interaction [2]. This poses a significant problem for the IoT network because most smart IoT applications require quick responses to customer requests, and the IoT network must meet many essential requirements for these applications, including latency, traffic reduction, privacy, and capacity[3].For instance, applications for healthcare monitoring should have excellent security and privacy for patient data, be quick to act in an emergency, and require a lot of bandwidth to communicate the patient's enormous amounts of data through the IoT network that is sensed every day[4]. By implementing the fog computing concept at the fog smart gateway, the smart monitoring system may have access to better systems and services including data mining, distributed memory, and warning service at the network edge [5].

Wireless Body Sensor Networks (WBSNs) typically consist of wireless sensor nodes that are small, lightweight, and positioned on, around, or on human bodies. These sensor nodes can monitor human body functions and environmental characteristics [6]. Recent advancements in applications for the Internet of Things (IoT) and remote health monitoring (eHealth) have boosted interest in WBSN. The variety of applications (healthcare, military, entertainment, etc.) demonstrates the necessity for more adaptable architectures and protocols in order to achieve higher performance.Monitoring of physiological and biochemical parameters including electromyography (EMG), electrocardiography, blood pressure, and blood sugar (ECG), and others may be accomplished with WBSN, a developing technology. When it was utilized in their line of work, WBSN originally got engaged in defense, but it has since evolved is currently utilized in many facets of daily life for the first time in human existence. Small, connected sensors that sense data from the patient's body make up the WBSN analyze it, and communicate with the medical server. These biosensors are developed with a specific goal in mind to fulfill the needs of end users [7]. for instance, to assess heart activity, an ECG sensor was developed. The aging population, high healthcare expenditures, and neonatal health are all urgent problems that need to be addressed right away. Medical facilities are effectively provided by doctors, but some illnesses such as cancer require ongoing patient monitoring; if therapy starts later, health improvement will be challenging, and patient mortality may follow.

The majority of chronic diseases, including cancer, obesity, cardiovascular disease, and diabetes, call for ongoing assessment of the patient's condition; however, this cannot be done simply relocating the patient. Patients who utilize a WBSN are physically more mobile and require shorter hospital stays [7,8,9]. Depending on how they have been used, sensor nodes can be categorized into three groups: implant nodes, which are inserted into the body's tissue or under the skin, body surface nodes, which are positioned 2 cm away from the patient's body, and external nodes, which are not in contact with the body but are a few centimeters to five meters away from it [10].WBSN consists of sensors positioned in, around or on the human body to monitor different physiological and biochemical parameters. These sensor nodes are capable of exchanging information with a network coordinator (sink), such as a smartphone or patient data aggregator (PDA), a robot, or another device that typically consumes less energy and has more processing power. It is in charge of providing the doctor with the patient's biological signal in real-time so that he may make the best judgments feasible [11,12]. The prevalent communication architecture of WBSN includes three levels: Intra-BAN communication, Inter-BAN communication and beyond BASN communication [13].

EEG biosensor is one of the wireless biosensors that commercially available in the markets [14]. EEG is a method for detecting electrical activity in the brain, frequently on the surface of the scalp. Ionic fluxes that are mediated by coordinated synaptic stimulation of brain neurons are what lead to these electrical events [15]. Depending on the application, EEG-based medical devices often need a varied number of sensors. For instance, whereas some healthcare systems only need a few electrodes, others need hundreds. Another factor affecting the number of sensors needed is the size of the skull, with toddlers needing fewer EEG electrodes than adults [16, 17]. The EEG system has anywhere from four to twenty-six electrodes [18]. These electrodes collect the brain's EEG signal, which is subsequently wirelessly transmitted to the coordinator (e.g., a PDA) for additional processing. Such apps frequently require large amounts of data to be recorded, sent, and analyzed. High-density EEG devices, for example, may sample at a rate of 1000 samples per second, with up to 100 electrodes, and each sample is encoded by two bytes.

As a result, the produced data rate from a single patient may reach 1.6 Mbps. Additionally, Medical data must be delivered to the Mobile-Health Cloud (MHC) every 10 seconds rather than every five minutes in situations requiring high-intensity monitoring [19,20].

1.2 Problem of Statement

In smart healthcare (s-health) systems, a massive amount of medical data is continuously recorded, processed, and communicated over the network, particularly in remote and constant monitoring applications. The rate of data generated from high-quality devices that consist of up to 100 electrodes can be 1.6 Mbps for every patient. High-intensive monitoring is required to be reported every ten seconds in emergency cases. Therefore, increasing in the transmission of health patient records over the network can increase the massive volume of data, bandwidth usage, errors generated during transmission, latency, and congestion over the IoT network.

1.3 Motivations

The main motivations to compress the EEG data at the fog gateway are:

- ❖ Reducing transmitted huge EEG data by removing the redundant data using compression approach.
- ❖ Reduce the exchanged data and long computation delay that have important impact on the speed of response of the IoT applications. It has the ability to improve the responsiveness of these applications.
- ❖ Save the network bandwidth through traffic reduction.

1.4 The Aim of Thesis

- Propose an Efficient EEG data compression approach at the fog gateway to improve the overall performance of the network .
- Integrate the Clustering approach with data compression to obtain a higher ratio of lossless compression on the EEG gathered data .
- obtain a higher ratio of lossless compression for the EEG gathered data.

1.5 Main Contributions of this Thesis

The main contributions of this work can be summarized as:

- A Lossless EEG Data Compression (LEDaC) technique for fog computing based IoMT network is proposed to reduce the EEG data at the fog gateway by compressing them before sending them to the cloud. The LEDaC exploit two efficient data reduction approaches: DBSCAN clustering and Huffman encoding.
- An Efficient Compression Technique (ECoT) for fog computing based IoMT network is proposed to reduce the EEG data at the fog gateway. Before transmitting them to the cloud platform. The proposed ECoT method combines two effective approaches: DBSCAN clustering followed by Delta encoding and Huffman-encoding at the gateway in the fog layer.
- A novel lossless EEG data compression (NoLEDaC) approach enabled by fog computing on Internet of Medical Things networks is proposed. To obtain a higher ratio of lossless compression on the EEG gathered data, the proposed NoLEDaC method combines two effective approaches: DBSCAN clustering followed by RLE and Huffman-encoding at the gateway in the

fog layer. The final compressed file is then sent to the cloud data center by the fog node.

1.6 Related Works

IoT Medical monitoring systems have significant challenges in compressing large amounts of patient EEG data while ensuring that the final destination receives accurate data. Lossless compression and lossy compression are the two forms of compression that exist. The biosensor-cloud interface can benefit greatly from lossless compression. It uses a coding scheme that effectively minimizes the number of bits required to represent specific data while ensuring that no data is lost during compression and decompression. On the other hand, Lossy reduction seeks to compress large volumes of data with the least amount of information loss possible. Various research on EEG compression techniques has been conducted, with various methodologies and algorithms presented.

In [21], Srinivasan et al. Real-time lossless EEG compression is proposed using an effective preprocessing method that arranges an electroencephalogram (EEG) signal in matrix form. An integer lifting wavelet transform serves as the decorrelator in the compression technique, and set partitioning in hierarchical trees serves as the source coder. In comparison to the traditional one-dimensional compression approach, experimental results reveal that the preprocessed EEG signal had better rate-distortion performance, especially at low bit rates, and shorter encoding latency.

In [22], Srinivasan et al.The authors investigate various electroencephalograph (EEG) signal lossless compression methods. they propose a straightforward computational pre-processing method that arranges the EEG signal into a 2-D matrix prior to compression. they discuss a two-stage coder with a lossy coding layer (SPIHT) and residual coding layer for compressing the EEG matrix (arithmetic coding). This coder is tuned to take full advantage of the source memory and the residual's i.i.d. characteristics. The JPEG2000 image compression standard, predictive coding-based reduction, and simple entropy coding are some of the other approaches they look into and compare to EEG compression. The Physiobank Motor/Mental Imagery database and the University of Bonn database are used to test the compression algorithms. In comparison to conventional vector-based compression, predictive coding, and entropy coding techniques, 2-D-based compression schemes produced higher levels of lossless compression. Pre-processing techniques led to a 6% gain in compression performance, and two-stage coding produced an additional 3% improvement.

In [23], Maazouz et al. propose a DCT-Based Algorithm for Multi-Channel Near-lossless EEG Compression. The electroencephalogram (EEG) signals are discussed in this article for compression purposes. In the temporal domain, The correlation in the EEG signal is used to compress the signal. Reduced transmission speed, energy utilization, and memory requirements are all benefits of data compression (reducing the cost accordingly). This article uses lossy compression based on the discrete cosine transform (DCT). This methodical process serves as the basis for the popular JPEG stills format. Information is lost during quantization. In addition to compression, entropy coding is used. Results are evaluated using the compression ratio and reconstruction quality specified in the PRD (percent root mean-square distortion). The suggested appeoach achieve suitable compression ratios for minimal distortion. Comparative to other works, they yields interesting CR with excellent reconstruction quality for values of PRD lower than 8%.

In [24], Birvinskas and Jusas propose Fast DCT algorithms for EEG data compression in embedded systems. This article discusses the use of fast Discrete Cosine Transform (DCT) algorithms for lossy EEG data reduction. This technique divides the signal into eight samples, each of which is DCT-transformed. The inverse transform is filled with zeros before transmission, and the least-significant transform coefficients are removed.It is concluded that when high speed and low processing complexity are needed, this approach may be used in real-time embedded systems. Depending on the data used, the reconstruction quality, measured in percent of root-mean-square (PRD) difference (PRD), ranges from 5% to 11% at a compression ratio of 4:1.

In [25], Karimu and Azadi propose a lossless EEG compression using DCT and Huffman Coding. Based on the characteristics of the DCT frequency spectrum and Huffman coding, a lossless hybrid compression technique for EEG is devised in this study. It generates DCT coefficients for EEG segments below 40 Hz (dominant components). The quantitative DCT parameters are then encoded using a Huffman encoder at the transmitter location. A zero set of DCT coefficients has been added above 40 Hz at the receiver location, then rebuild EEG segments using inverted DCT. The technique has been used on the University of Bone database's five groups (labeled A-E). The research shows that this approach can increase these sets' average compression ratios by 1.78, 1.94, 2.66, 3.35, and 1.78 times, respectively, above the best results found in the literature.

In [26], Alsenwi et al. suggest a performance analysis of hybrid lossy/lossless compression techniques for EEG data. Two lossless compression methods are utilized in this study, discrete cosine transformation (DCT) and discrete wave transformation were used to turn random EEG data into high frequency data (DWT). In order to achieve a high compression ratio without experiencing signal distortion, it is wise to use a lossless compression approach following lossy compression.The Run Length Encoding (RLE) and Arithmetic Encoding lossless compression techniques have been used. The overall compaction and rebuilding durations (T), compression ratio (CR), root mean square error (RMSE), and structural similarity index (SSIM) are analyzed to determine the effectiveness of the proposed method.

In [27], A. Fathi and B. Hejrati, A new lossless compression approach that is both efficient and easy has been presented. Finding the connection between and within channels is referred to as "correlation" in this context. A preprocessing step to extract intra-channel correlation is included in the first stage of the differential pulse code modulation approach. The channels are divided into discrete groups, and each group's centroid is identified and encoded using arithmetic coding. The difference between the centroid and other channel data is identified and encoded in the second stage using arithmetic coding.

In [28], Saeed et al. propose a hybrid compression technique with data segmentation for electroencephalography data. Electroencephalogram (EEG) data volumes are enormous due to the prolonged recording length, high sampling rate, and numerous electrodes utilized in medical applications. In order to effectively convey and preserve data, greater room and bandwidth are needed. In order to transmit EEG data effectively using less bandwidth and store it in a smaller amount of space, EEG data compression is a crucial issue. In this study, an efficient EEG compression method is provided. The discrete cosine transform is then used to process the EEG data once it has been divided into N segments (DCT). Any values below the threshold are set to zero after the modified parameters go through a threshold operation. The acquired parameters are then encoded using the Run Length Encoding (RLE) method.

The EEG signal can be retrieved using a reverse approach. To determine if the suggested technique is effective, the total compression and reconstruction time (T), compression ratio (CR), and root mean error difference (PRD) ratio are determined. According to simulation findings, data segmentation significantly reduces compression time.

In [29], J.Azar et al. A lossless compression approach based on the Discrete Wavelet Transform (DWT) and a polynomial interpolation lifting mechanism have been proposed . Prior to transmission to the coordinator, the number of bits used to hold detected medical data is decreased to save energy and extend the life of the health network. There are no additional mathematical computations or memory requirements with this strategy. When the results of the compression technique were compared to those of other compression methods, including the lifting scheme-based Haar transform, the discrete cosine transform, and simple delta encoding, they discovered that theirs had a greater compression ratio. On the other hand, Lossy reduction aims to compress a huge quantity of data with the least information loss.

In [30], AlNassrawy et al. a potential EEG Fractals compression model is suggested for decreasing the sent EEG traffic from Patient Data Aggregator (PDA) to the destination (doctor, smart hospitals, emergency response, etc.). The suggested model helps with EEG patient data transmission and enhances Wireless Body Sensor Network by reducing network traffic. The determined Fractals Block Size is discovered to play a crucial part in providing a greater Compression Ratio (CR) and driving the necessary Percentage Residual Difference as the primary model metrics are examined (PRD). The proposed model completely outperformed the other strategies when results and performance were examined. The resulting CR can be as high as 160 while maintaining a PRD below 1.

In [31], M. Pushpalatha and P. Rajaseker have successfully sent EEG data by using a lossless compression technique called the Huffman-based discrete cosine transform. The discrete cosine transform and inverse discrete cosine transform are proposed to improve data privacy and decrease data complexity. The primary objective of this research is to achieve a high accuracy ratio when reconstructing the original data after compression and transmission without any losses in the shortest amount of time possible. To minimize noise and send the original data, preprocessing and sampling are done from the beginning. When compared to existing techniques in various data transformations, the discrete cosine transform-based Huffman quantization method achieves high performance metrics in terms of peak signal-to-noise ratio, quality score, and compression ratio.

In [32], Samiran Das and Chirag Kyal presented an effective tensor technique for data compression reduction. The best-compressed tensor size is determined by recommending that the multichannel EEG data be reshaped as a tensor first. Then use tensor decomposition to build a core tensor that is substantially smaller but still retains all of the tensor's information. To achieve such a high compression ratio, the process' core tensor retains every bit of data from the original signal. On the decoder side, they rebuild the original tensor by transforming the core tensor into factor matrices in the transposed order. They investigated the proposed compression technique using publicly accessible statistics and performance metrics. In prior research, they determined that their compression approach surpassed state-of-the-art techniques.

In [33], S.k. Idrees and A.k.Idrees, proposed New fog computing enabled lossless EEG data compression scheme in IoT networks. To reduce the quantity of IoT EEG data transferred to the cloud, a new fog computing enabled lossless EEG data compression technique is described in this study. Clustering and compression are the two steps of the EEG data compression technique. Using agglomerative hierarchical clustering, the incoming data is first grouped into clusters. In the second phase, the Huffman encoding is applied to each cluster that has been created. Finally, smaller clusters' compressed data are concatenated and uploaded from fog to cloud. Several tests have been carried out, and the results demonstrate that using the recommended combined Hierarchical Clustering and Huffman Encoding (HCHE) approach reduces the amount of EEG data transferred to the cloud platform substantially. The suggested HCHE approach produces an average compression power of (4.33), which is more than double that of certain current methods.

In [34], Giuseppe Campobello et al. proposed a novel near-lossless compression technique for compressing electroencephalograph (EEG) data effectively. Particularly, their proposed method aims to give a low-complexity solution that may be used in EEG monitoring devices that are worn on the body. The method is based on a straightforward and energy-efficient encoding technique that may be implemented on microcontrollers with low power consumption. Despite its simplicity, the suggested method yields compression ratios comparable to those of more complex state-of-the-art algorithms, as demonstrated by comparisons with real-world data.

In [35], Al-nassrawy et al. proposed a novel lossless EEG compression model using fractal combined with fixed-length encoding technique. This research presented a lossless Fractals compression method for reducing the amount of EEG data transferred from the patient data aggregator (PDA), the gateway cloud reducing the amount of data traffic passing through the network, the proposed approach improves data communication in WBSNs. This method is tested and compared to various existing approaches, with the findings indicating that the proposed method outperforms the others.

In [36],A.K.Idrees et al, was proposed an IoMT network using edge-fog computing to enable lossless EEG data compression and epileptic seizure detection, The recommended strategy accomplishes three goals. At the Edge Gateway, lossless EEG data compression based on a hybrid approach of k-means Clustering and Huffman Encoding (KCHE) is used to minimize the quantity of data transported from the Edge to the Fog gateway. The Epileptic Seizure Detector-based Naive Bayes (ESDNB) approach is used to identify the patient's epileptic seizure state at the fog gateway. Thirdly, it reduces the amount of IoMT EEG data transferred to the Cloud using the same lossless compression technique as in the first phase. Numerous experiments were conducted to demonstrate the proposed strategy's efficacy, and the comparison outcomes showed that the KCHE lowers the amount of EEG data supplied to the fog and Cloud platform while still offering acceptable epileptic seizure detection. The average compression power of the proposed KCHE is four times more than that of the existing EEG data compression techniques (Z, F, N, O, S). Additionally, the suggested ESDNB exceeds earlier techniques in terms of accuracy, attaining 99.53 percent to 99.99 percent precision using the dataset from the University of Bonn.

Table 1.1: Summary of the Related works

No. Ref	name & year	Lossless	Lossy	Data Set		Technique
				Bonn	Others	
[21]	Srinivasan et al(2010)	✓		✓		Lifting wavelet transform.
[22]	Srinivasan et al(2011)	✓		✓	✓	Shift+arithmetic coding.
[23]	Maazouz et al(2015)		✓		✓	DCT algorithm
[24]	Darius Birvinskas&V. Jusas(2015)		✓		✓	Fast DCT Algorithm
[25]	R Ykarimu &sAzadi (2016)	✓		✓		DCT + huffman coding
[26]	Madyan Alsenwi et al(2016)	✓	✓		✓	DCT + DWT (lossy) RLE + arithmetic encoding (lossless)
[27]	B. Hejrati & A. Fathi (2017)	✓			✓	Arithmetic coding
[28]	Mohammed saeed et al(2017)	✓	✓		✓	DCT + RLE
[29]	J.Azar et al (2018)	✓			✓	DCT lifting
[30]	AlNassrawy et al(2020)		✓	✓		Fractal compression technique
[31]	p.Rajasekar&M.Pushpalatha(2020)	✓			✓	Huffman encoding
[32]	Samiran Das & Chirag Kyal(2021)		✓		✓	Tensor truncation method
[33]	Sara.k.idrees&Ali.k.idrees(2021)	✓		✓		Hierarchical clustering + Huffman encoding
[34]	Giuseppe Campobello et al(2021)	✓			✓	A simple and efficient

						encoding scheme
[35]	Al-nassrawy et al(2022)	✓		✓		Fractal compression with fixed-length encoding
[36]	A.K.Idrees et al (2022)	✓		✓		k-means Clustering +Huffman Encoding

1.7 Thesis Outline

The rest of thesis has the following arrangement :

Chapter Two: declares the essential background information about Wireless Body Sensor Network (WBSN) in terms of main structure, communication architecture, hardware devices types, characteristics, medical applications, challenges and fog computing with its features. It also offers EEG sensor and data clustering and compression approach.

Chapter Three: explains the proposed techniques,in addition to the suggested method's algorithms for EEG compression and decompression.

Chapter Four: discusses and demonstrates the practical effort to evaluate the suggested methodologies, and in this chapter, the performance evaluation is compared to other comparable earlier work.

Chapter Five: present the conclusions and future works.

Chapter Two

Theoretical Background

2.1 Introduction

One of the most efficient communication protocols of the twenty-first century is the Internet of Things. All items in everyday life become part of the Internet in the IoT ecosystem due to their connectivity and processing capability. IoT expands and popularizes the notion of the Internet. The Internet of Things (IoT) provides seamless communication across a wide range of electronic devices [37,38]. As a result, IoTs in some areas, such as healthcare technology, have become more imaginative. IoTs in healthcare entailed a range of low-cost sensors, both wearable and implantable, that enabled older people to get current medical healthcare services from anywhere, at any time, and enhanced their quality of life [39].

These devices must communicate their health data to an external server, and utilizing a cable connection for this purpose is too expensive in terms of deployment and conservation; consequently, using a wireless interface simplifies and reduces the cost of the application. WBSN allows the patient to move around and eliminates the requirement for the patient to stay in the hospital for a long time [40].

A significant amount of data is acquired, evaluated, and transmitted in health apps due to ongoing monitoring and real-time feedback for users [40]. As a result, reducing redundancies in medical data must be employed as a reduction approach to enhance health networks and conserve energy.

2.2 IoMT Architecture

An IoMT system consists of three basic operational layers [41]:

I. **Data collection layer:** This layer is responsible for the collection of medical data from various sensor devices attached to the patient/test subject that needs to be monitored/examined [41].

II. **Data storage layer:** This layer is responsible for storage of big data collected from various sensors and transmitted through the Internet [41].

III. **Data processing layer:** This layer analyzes the data stored in servers to generate the required response through application of computing algorithms. Also, the compilation and visualization of the results are done here [41].

2.3 WBSN architecture

WBSNs are made up of three distinct components, regardless of whether the health network is static or dynamic: wearable sensor nodes, the health network coordinator, and the destination node [42].

A WBSN is built by merging these components. There are also internet-connected patient monitoring and servicing systems, which may be accessed remotely by doctors or smart hospitals. The WBSN Architecture is depicted in Figure (2.1) [43].

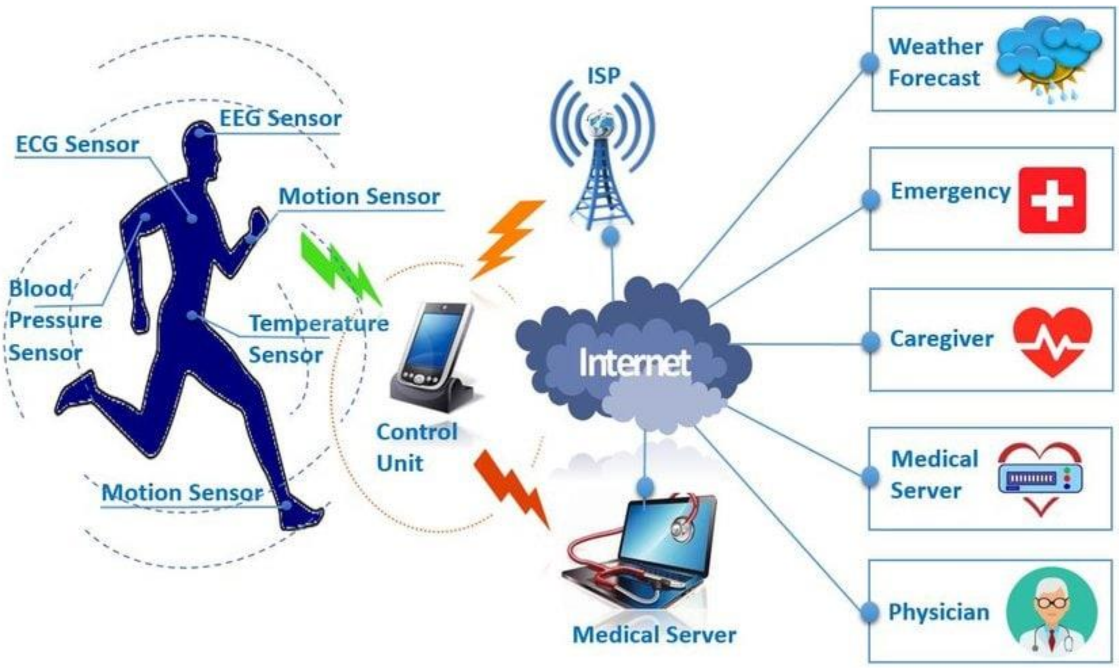


Figure 2.1: Wireless Body Sensor Network Architecture[43].

2.3.1 Wearable Medical Sensors (WMS)

The initial component of WBS is a wearable system that should be proactive and affordable. They are put with care on the body. They exchange data via the ZigBee protocol (IEEE 802.15.4). Their goal is straightforward: acquire relevant data without hindrance. Sensors (such as ECG, accelerometer, and others) are usually combined into a single item, such as suits, to simplify problems. A master node (smart node), also known as a micro gateway, is chosen to gather and transfer data from all sensors to the NC[42].

Commercially available wireless body sensors include ECG sensors, blood pressure sensors, CO2 gas sensors, humidity and temperature sensors, blood glucose sensors, EEG sensors, and pulse oximetry [14].The current thesis, the concentration is placed on EEG wireless body sensors.

2.3.2 Health Network Coordinator (HNC)

The second component of WBSNs is a multi-hop device that collects health data from biosensors and sends it to the server. It has significantly greater processing capabilities and power than body sensors. Its positions are [42]:

- ❖ Overseeing and managing WBSN nodes.
- ❖ Collecting and analyzing health data from many types of sensors.
- ❖ The interface for the leading system is being prepared.
- ❖ Securely communicating with a medical service provider.

2.3.3 Base Station (BS)

The most appealing of the three WBSN components, with the quickest communication rate, the greatest battery power, and the most processing capabilities, acquires data from the network coordinator. The bandwidth of the base stations, such as the Star Gateway Net-bridge base station [42], must be considerable in this instance.

2.4 Fundamentals of EEG

The electroencephalogram (EEG) is a technique for monitoring electrical activity in the brain, often along the scalp's surface. These electrical events are caused by ionic fluxes mediated by coordinated synaptic stimulation of brain neurons [15], and they appear as rhythmic voltage fluctuations with amplitudes ranging from 5 to 100 V and frequencies ranging from 0.5 to 40 Hz. Examining the important frequencies and amplitudes of EEG waves in various parts of the brain may offer information about a person's medical or mental condition [44]. Brain waves are classified into five frequency bands based on their frequency: In general, the delta waveform (14 Hz) is the slowest and has the greatest amplitude. The delta band may be apparent in newborns and adults during deep sleep.

• Children, fatigued adults, and those reliving memories all have theta (4.88 Hz) frequency. Theta waves have an amplitude of less than 100 V in most cases.

• Active contemplation, concentration, and focused attention are all connected to beta (12.25 Hz). Additionally, doing or watching others conduct physical movements boosts Beta power. Beta waves have an amplitude of less than 30 volts in most cases.

• During multimodal sensory processing, gamma (above 25 Hz) is detected. The amplitude of gamma patterns is the smallest.

As a result, while investigating brain waves, it is critical to pay attention not only to the dominant frequency but also to the recording from a specific portion of the brain.

In order to collect EEG data, electrodes with low impedance are used. Electrodes are a kind of electrode. It can be applied with conductive gel, as a "wet electrode," or straight to the skin, as a "dry electrode." One example of a set of guidelines for positioning and identifying electrodes on the scalp is the 10-20 standard [45]. The electrodes in the (10-20) standard is located along latitude and longitude and are classified according to the lobe to which they belong. The poles in the right hemisphere, on the other hand, may be allocated even numbers, while the poles in the left hemisphere may be assigned odd numbers (as shown in Figure 2.2 [46]).

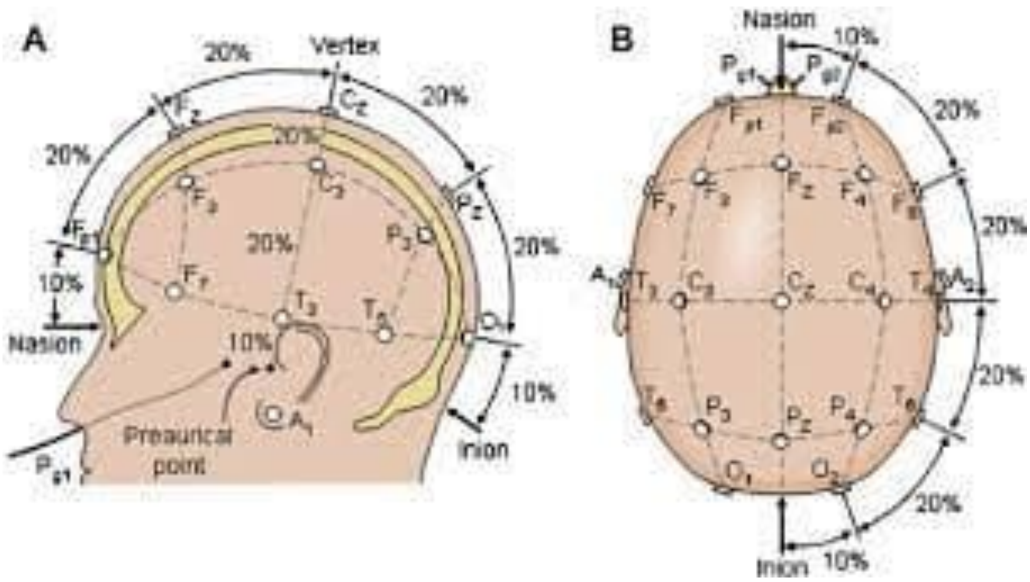


Figure 2.2: The placement of the Electrode (A) Left (B) above the Head in the 10-20 Standard[46].

2.5 WBSN Communication Architecture

When looking at the entire WBANs ecosystem, data transmission may be separated into many phases. It should be noted that when a person moves in this circumstance, their complete body may also move. As a result, in this scenario, the sensor's positioning may change, exhibiting the dynamic nature of WBANs. The WBANs standard recognizes three layers of communication [47].

2.5.1 Tier-1: Intra-BAN communication

This grade supports both wired and wireless connections. Zimmerman [48] suggests this method of communication. Only the sensors and the sink are linked in intra-BAN communication. This tier's communication range is roughly 2 meters within and around the human body. Because the sensors are largely positioned within this connection range, this layer is vital. As a result, communication is restricted. Communication technologies such as Bluetooth and ZigBee are used in this layer. Sensors inside this layer collect data on physiological features and deliver it to a sink. The sink is in responsible of data processing and transmission to Tier 2 [47].

2.5.2 Tier-2: Inter- BAN Communication

This layer handles communication between the sink and one or more access points. Another possibility is that access points are developed by infrastructure or deliberately placed in dynamic circumstances to handle emergency situations effectively. The purpose of this layer is to connect numerous easily accessible networks, such as WBANs and mobile phone networks (or the Internet). Cellular, 3G/4G, ZigBee, Bluetooth, Wireless Local Area Networks (WLANs), might all be included at this layer [47].

2.5.3 Tier-3: Beyond BASN communication

Metropolitan Area Networks inspire this layer (MANs). The medical sensor is linked to the Internet or another network, which allows data to be delivered to receivers, allowing doctors and health specialists access to the information. The present might be given to a doctor or a nurse. The information might be saved in the patient's database. Tier-3 is so heavily reliant on the database. The database includes information on the patient's profile, users, and medical history. When this occurs, the doctor is notified that the patient's health is deteriorating and may take necessary action before the patient arrives at the hospital. The medical environment and database, which preserve the user's medical history and profile, are the most significant components of Tier-3s. As a result, physicians and patients can be notified about medical crises by text message or the Internet. Tier 3 also ensures the recovery of any critical patient data that may be required for treatment. Tier-1 sink-ins may communicate with an AP through 3G/4G/GPRS instead, depending on the application [47].

2.6 Hardware Devices Types

There are generally three heterogeneous devices in wireless health network can be categorized according to their functionality [49].

i. **Wireless sensor node:** It collects data on a physical stimulus and, if necessary, processes and reports the data via wireless transmission media. It consists of a number of components, including sensor hardware, a processor, memory, a power supply, and a transmitter/transceiver[49].

ii. **Wireless actuator node:** The actuator nodes respond to data received from a sensor. Actuator components are equivalent to sensor components: actuator hardware (for example, medicine control hardware, which may include a medicine bank), power generator, CPU, memory, transceiver, or receiver[49].

iii. **Wireless personal device (PD):** An external portal, monitor, actuator, or display/LEDs support device on the screen stores and sends all of the information acquired by the actuators and sensors. The primary components are a huge processor unit, a memory unit, a power unit, and a transceiver [49].

2.7 Characteristics of Sensors

Sensor nodes in WBSN have a variety of properties that make them suitable for use in a variety of emerging applications [14]. Several characteristics can be summarized as follows:

2.7.1 Energy Efficient

These nodes are particularly energy efficient due to their architecture. Energy resources are managed by energy management solutions that optimise and extend the life of the WBSN[14].

2.7.2 Heterogeneous

The WBSN employs a variety of sensors, each of which performs a specific job; for example, some nodes are dedicated to sensing temperature, while others are dedicated to measuring brain activity. As a result, processing, memory capacity, and power consumption vary between sensor nodes[14].

2.7.3 Cost Effective

Due to the fact that sensor nodes in WBSN operate for extended periods of time due to their efficient power consumption and deployment on the body region (small area), only a limited number of nodes are required for network building and replacement operations. This lowers the total cost of network construction[14].

2.7.4 Simple

WBSN sensor nodes are extremely lightweight and small in size, making them easy to wear and transport from one location to another[14].

2.8 WBSN Medical Applications

These applications cover a wide range of fields, including health care, aid to the elderly, and emergency response. This section will provide an outline of potential medical applications [50].

2.8.1 Telemedicine and Remote Patient Monitoring:

Through the use of integrated health information systems and telecommunications technology, telemedicine [51] is essential in the distant delivery of patient care and enables medical practitioners globally to treat a larger patient population. Upon receiving health information, medical workers might begin therapy beforehand [52].

2.8.2 Cancer Detection:

Cancer was recently revealed to be the deadliest sickness on the globe, spreading at such a quick rate that many people have perished as a result of late diagnosis. Nitric oxide [52] can be detected using a sensor at the probable spot. These sensors can tell the difference between dangerous and non-cancerous cells, allowing doctors to detect cancer at an early stage [50].

2.8.3 Rehabilitation:

Rehabilitation is the process of resuming a person's normal activities following hospitalization. Throughout rehabilitation, patients must be closely followed to ensure they can resume their usual lives [50].

2.8.4 Glucose Level Monitoring:

Additionally, one of the most common diseases, particularly among persons in the workforce, is diabetes. Further complicated illnesses including heart disease, stroke, hypertension, and blindness may follow from this [50].

2.8.5 Biofeedback:

Using WBSNs and sensor data, it is now feasible to self-monitor the human body. The patient gets sensors implanted within or on their body so that they are immediately aware of any health concerns and do not need to visit the doctor for routine examinations. The patient can do numerous daily regular tests, such as electrocardiography (ECG), temperature analysis, and blood pressure detection, on his or her own [53].

2.9 WBSN Challenges

The size of the WBSN node is the primary requirement. The small size of the sensor node limits the size of the battery, ensuring that the system as a whole is energy-efficient. As a result, the design of the WBSN faces a number of difficulties [50], some of which are stated below:

2.9.1 Accuracy

Reliability recommends high levels of accuracy; thus, the sensors should provide accurate and fast patient health information. If there are any time or information delays, there may be a risk to human life or a delay in treatment. As a result, the WBSN should provide a high level of accuracy, resulting in a trustworthy network [50].

2.9.2 Information Awareness

The sensors' behavior must adapt to the user's conditions and the surroundings, such as the weather outside, the time of day, and so on. This situation should not have an influence on the WBSN's operation or the supply of erroneous patient health information [50].

2.9.3 Comfort and Efficient Response

One of the most critical problems is to make WBSN user-friendly and enjoyable for patients, particularly the elderly. As a result, the application should be simple to use and small in size. Patients may have trouble using a device because of its size, difficulty of construction, or other causes [50].

2.9.4 Energy Consumption

The sensor is powered by a small integrated battery [51]. They quit working as soon as the battery dies. As a result, data transmission has been proven to consume more energy than processing [50].

2.9.5 Fault Tolerance

Intruder assaults, sensor device failure, and bad environmental conditions are all common problems in sensor networks. Even if some nodes refuse to join, the network must nevertheless provide accurate and trustworthy services. As a result, information about the human body may need to be collected by more than one node [50].

2.9.6 Privacy and Security

Patient health information must be protected when transferred through wireless media. As a result, the data can be encrypted before transmission using a number of methods [54].

2.9.7 Topology

The sensors are embedded in or cover the human body. The movement of the patient influences the route architecture [55]. These networks should be able to tolerate frequent changes in network topology and continue to function even as the topology changes.

2.10 Fog Computing

Fog computing is a paradigm or platform that provides a few restricted functionalities, such as networking, processing, and storage, but in a more scattered way than typical cloud computing [56]. The platform provides a trustworthy solution for Internet of Things (IoT) apps and devices, although it is prone to delays. The term "fog computing" was coined by Cisco, while other firms and researchers have described it from other perspectives [57]. Figure 2.3 depicts the fog's computing architecture.

2.10.1 Features of Fog Computing

Fog computing is a cloud computing advancement that is more analogous to IoT-capable devices. Fog computing acts as a bridge between edge devices and CC, bringing processing, networking, and storage capabilities closer to edge devices [58]. The key properties of fog computing are as follows:

a. **Adaptability**: The surrounding environment is monitored by a vast network of sensors. The fog provides distributed processing and storage capabilities that can work with such a wide range of end devices.
b. **Real-time communications**: When using cloud batch analysis, fog computing solicitations offer simultaneous communication between fog nodes.
c. **Physical distribution**: In contrast to the integrated cloud, fog offers decentralized applications and services that may be hosted anywhere.
d. **Less latency and position awareness**: Since fog is close to edge devices, there is reduced waiting time for processing edge devices' information. Additionally, it helps with position responsiveness by allowing fog nodes to be hosted everywhere.
e. **Compatibility**: Through a variety of service providers, fog modules can adapt to and work with many platforms that are not the same.
f. **Provisions for web-based analytics and integration with cloud**: In order to play a crucial role in the speed and computation of the information close to the edge devices, the fog is placed between edge devices and cloud.
g. **Heterogeneity**: Since they are created by different firms and have different origins, edge devices or fog nodes require hosting according to where they will be used. Fog may therefore adapt to different systems.
h. **Provision for flexibility**: The ability of fog solicitations to connect directly to devices like mobiles and so enabling flexibility approaches, such Locator ID Separation Protocol (LISP), which

requires a distributed indexed system, is one of its key characteristics.

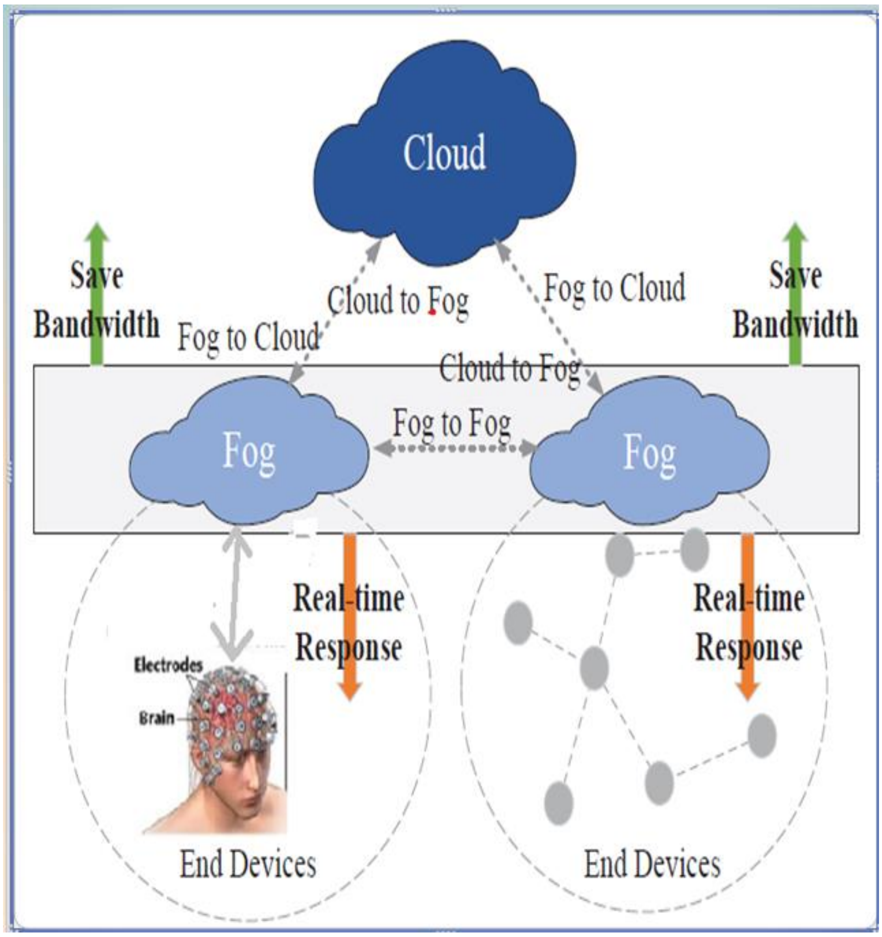


Figure 2.3: Fog Computing Architecture.

2.10.2 Fog Computing Advantage

Fog computing is a network-edge implementation of the cloud computing idea. Despite the fact that fog computing employs the same resources as the cloud (computing, network, and storage), as well as many of the same features and methodologies, it provides significant benefits for IoT-powered devices (multi-tenancy and virtualization). One of the benefits of fog computing is increased business agility. With the right approaches, fog computing applications may be designed and deployed quickly. Fog computing applications may also analyze device needs and configure them to work properly. Because of its low latency, fog computing can support the functionality and core of real-time applications such as gaming and video streaming [56]. In terms of scalability and geographic distribution, fog computing provides storage and compute to widely scattered applications. Fog computing decreases network bandwidth and lowers operating costs. Flexibility and heterogeneity are other features of fog computing. For example, it facilitates cooperation for unique services across various infrastructures and physical environments. The proximity of fog computing to end devices allows for easier scalability of services and associated devices. Finally, it is well-known for its scalability.

2.11 Data Clustering

Data Clustering is an unsupervised learning strategy that groups data into relevant subclasses so that items within a subclass are more similar to one another than objects within other subclasses. Clustering techniques are used in many areas, including image analysis, pattern recognition, knowledge discovery, and bioinformatics. The problem of locating clusters with varied shapes and determining the input parameters of algorithms with minimal domain knowledge requirements and acceptable performance on large datasets when using clustering techniques in geographical databases [59].

2.11.1 Types of Clustering Algorithms

Clustering can be done for any reason. A specialized clustering approach is designed for a specific set of applications. Each algorithm has its own way of describing the proximity of data points. More than 100 clustering algorithms have been proposed and explored to date. All of them may be divided into five subcategories [60]. Figure 2.4 clearly depicts the categorization of clustering approaches [61].

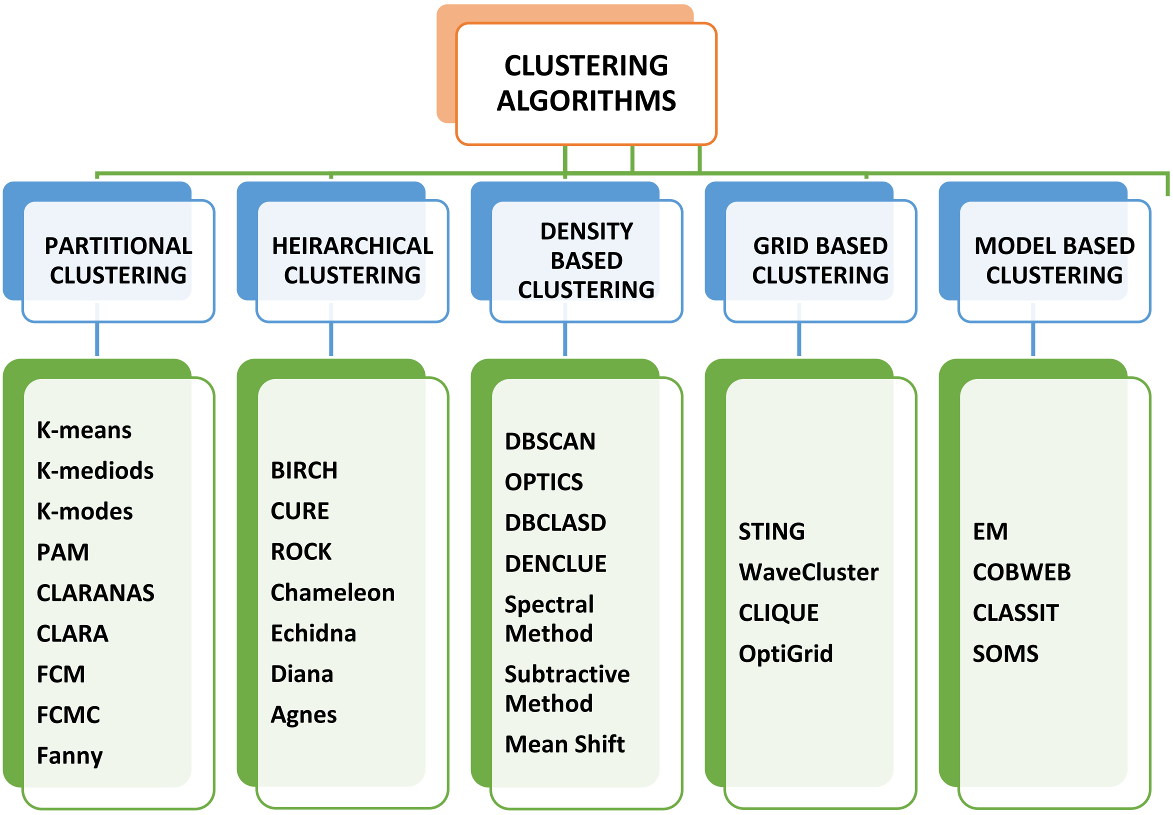


Figure 2.4 : Categorization of Clustering Algorithms

2.11.2 DBSCAN Clustering

The first density-based clustering algorithm is DBSCAN. It was created to cluster data of any shape in high-dimensional spatial and non-spatial databases and was first proposed by Ester et al. in 1996. The key idea of DBSCAN is that each object in a cluster must have a neighborhood within a set radius (Eps) that contains at least a specified number of objects (MinPts), which implies that the neighborhood's cardinality must be greater than a predetermined value. The ϵ-neighborhood of an arbitrary point 'p' is defined as,

$$N_{Eps}=\{\text{q c D / dist(p,q)} < \text{Eps}\}$$

Here, D is the database of objects. If the c –neighborhoods of a point P at least contain a minimal number of points, and then this point is called core point. The core point is defined as:

$$N_{Eps}(\text{ P}) > \text{MinPts}$$

Here Eps and MinPts are the user's specified parameters which mean the radius of the neighborhood and the minimum number of points in the c-neighborhood of a core point respectively. If this condition is not satisfied, then this point is considered a non-core point [62].

DBSACN searches for the clusters by checking the ϵ-neighborhood of each object in the dataset. If the ϵ-neighborhood of an object p contains more than *MinPts*, a new cluster with p as a core object is created. It then iteratively collects directly density-reachable objects from these core objects, which may involve the merge of a new density- reachable cluster. The process terminates when no new object can be added to any cluster [63].

2.11.3 The Strength of DBSCAN Algorithm

DBSCAN algorithm has following advantage[64]:

- In the case of DBSCAN clustering, it is not required to determine the number of clusters in the data at the beginning.
- It requires only two parameters and does not depend on the ordering of the points in the database.
- It can identify clusters of arbitrary shape. It is also able to identify the clusters completely surrounded by different clusters.
- DBSCAN is robust to outliers and has a notion of noise.

2.12 Data Compression

Data compression is the process of converting a collection of text, audio, and video files into a smaller-sized database that may be used to recover the original file without losing any important data. Because the compressed file requires less storage space and bandwidth to communicate, this approach provides for more efficient data storage and transit. When a collection of files in various formats necessitates a large amount of space, it is advised that a folder or database be converted into zip format to decrease its size and prevent data loss during transmission. ZIP files are a good example of file compression. Only compression technologies are capable of accomplishing this [65].

Data compression includes compressing data to a smaller size than the original data in order to save storage space and speed up network transfers. Data compression is both achievable and effective since duplicate data form a large fraction of all data in the actual world. These approaches can be lossless (recovering all of the original data in its original format) or lossy (some original data bits cannot be restored upon decompression). The procedures used to restore the original data are known as decompression algorithms. The primary categories of data compression are given below.

Because of the development of various data compression methods, it is required to analyze the approaches and choose the algorithm to use in a specific situation. There are four types of data compression methods. It is depicted in Figure (2.6) below [66]:

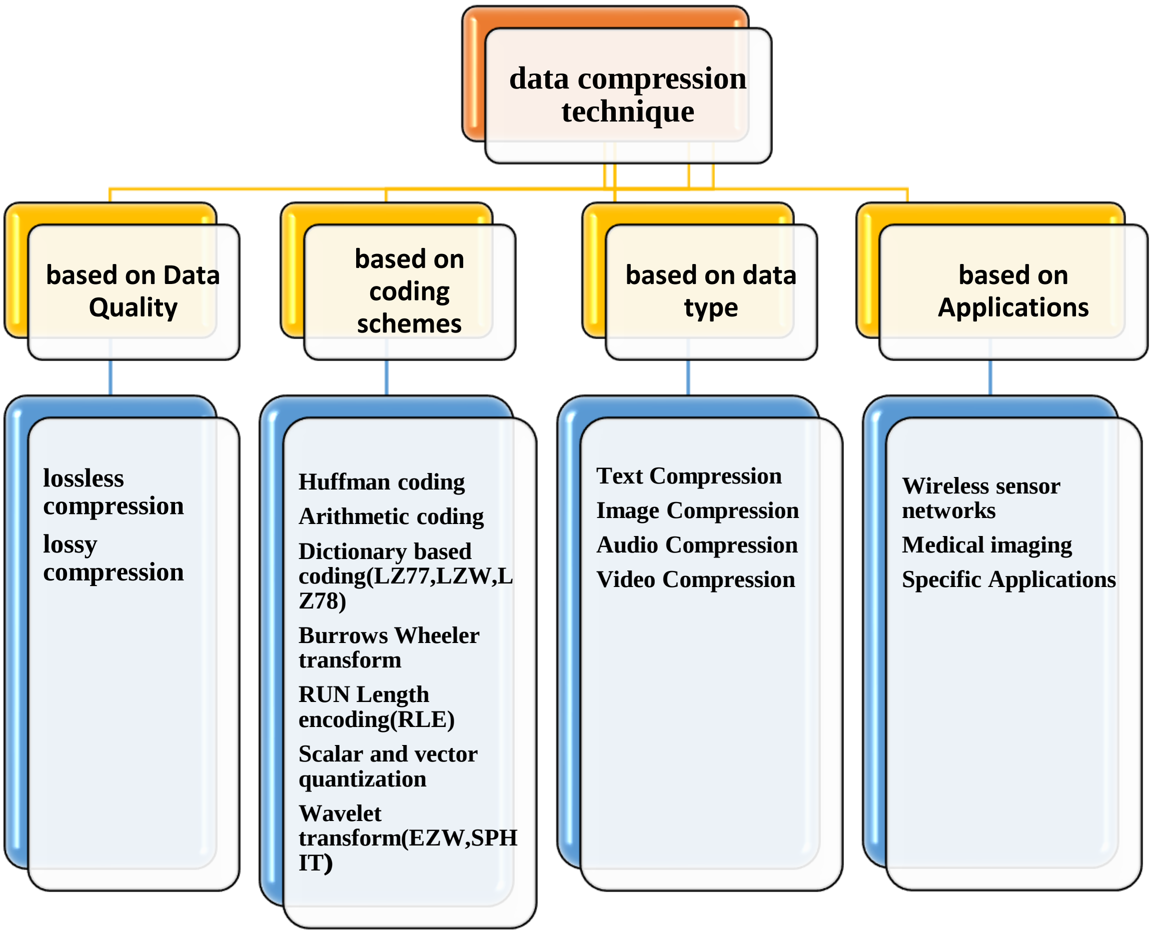


Figure 2.6: Classification of Data Compression Techniques.

2.12.1 Lossless Compression

When a file is compressed, all of the original data bits are preserved and may be recovered via file decompression. The data is only effectively stored in its compressed form, as the name indicates, but nothing is lost. Data must be restored without loss for mission-critical tasks [65].

2.12.2 Lossy Compression

It signifies that some data may be lost during the decompression process. This compression method is based on the assumption that modern data files contain more information than the ordinary human can grasp. As a consequence, unnecessary information can be deleted [65].

In the current thesis, lossless compression methods will be focused on.

2.12.3 Run-length Encoding

RLE is a form of lossless compression. The goal of RLE is to replace a data value's subsequent repeated occurrences with only one occurrence followed by the number of occurrences. This is particularly useful when working with data including a high number of such runs [67, 68].

2.12.4 Delta Encoding

Delta encoding, commonly known as relative encoding, is a basic lossless compression method that operates by calculating the difference between subsequent samples [69]

2.12.5 Huffman Coding Algorithm

Huffman coding achieves compression rates ranging from 20 to 90 percent. This compression method entails bit data encoding in binary bits. Following the encoding technique, the data may be decrypted using Huffman tree tracking from the tree's root using provided sequences. The key advantage is that it is simple to use in terms of computing. It shortens the standard code for representing alphabetic letters. It substituted each character with variable-length code based on the character's relative frequency [31].

Algorithm 2.1: Huffman Compression Algorithm

Input: F

***Output**: F^C: compressed file*

```
F^C ← QU();
for i ← 1 to Length(F ) do
    ND ← Node{F };
    F^C.PUSH(ND);
end
while F^C.Size() ≠ 1 do
    ND ← NewNode();
    ND.Left ← X ← F^C.POP ();
    ND.Right ← Y ← F^C.POP ();
    Frequency(ND) ←Frequency(XD) +Frequency(YD);
    F^C.PUSH(ND);
end
return F^C ;
```

Algorithm 2.1: Huffman Decompression Algorithm

***Input**: HTR : Huffman tree root, BS : the bit stream is needed to be decompressed.*

***Output**: F^D : decompressed file*

L ← Length(BS);
for** j ← 1 **to** L **do
F^D ← HTR;
while** F^D.LEFT /= NULL and F^D.RIGHT /= NULL **do
if** (BS_j = 0) **then
F^D ← F^D.LEFT;
end
else
F^D ← F^D.RIGHT;
end
j ← j + 1;
end
end
return F^D ;

2.13 EEG Datasets types

This section summarizes the most frequently used publicly available datasets for EEG data analytics.

2.13.1 Motor Motion Array/EEG Image Data from Physionet

The Physionet EEG Motor Movement/Imagery dataset is one of the most well-known and readily available datasets developed by the BCI2000 [70] technology. This dataset includes EEG recordings from 109 healthy participants who participated in a range of motor and imaging tasks. 64 channels of EEG waves were recorded for three 2-minute rounds for each of the four activities done by each participant at a sampling rate of 160 Hz for two 1-minute baseline runs (one with open eyes, the other with closed eyes). Consider opening and closing your left or right hand, as well as your fists and feet. Visualize opening and closing both feet [71].

2.13.2 BCI Competition II

This huge dataset comprises a comprehensive record of real-world BCI performance from three trained people across ten sessions. In each trial, the patient sat in a reclining chair in front of a video screen and was asked to remain motionless throughout the performance. On the scalp, 64 EEG channels were recorded, each relating to an electrode on the right ear (amplification 20000; band-pass 0.1–60 Hz). All 64 channels were digitized and stored at 160 Hz. Cursor movement was controlled online using only a few channels [72].

2.13.3 BCI Competition III Dataset

The datasets are made up of EEG recordings of one to five persons engaging in various activities. EEG recordings of three subjects (designated as K3b, K6b, and L1b) were made as part of one of these datasets utilizing the Neuroscan amplifier with 62 EEG channels (60 electrodes + 2 reference electrodes) and a sampling rate of 250Hz. Participants in the study imagined four bodily parts moving in response to this. There are triggers on the left, right, foot, and tongue. These data sets are being used in a variety of academic initiatives [73].

2.13.4 BCI Competition IV Datasets

To continue the BCI contests, the fourth BCI competition, involving five independent datasets, was organized in 2008. These datasets feature more subjects and tasks. In Dataset 2a[74], for example, the EEG signals of nine persons were recorded at a sampling rate of 250 Hz on 22 EEG channels and 3 EOG channels. Four distinct activities were assigned to the participants, including imagining left- and right-hand motions and both legs days. Each session is made up of six runs separated by short pauses. These datasets have also been utilized in other studies [75].

2.13.5 EEG Database at UCI KDD

This dataset was taken with 64 electrodes and a sampling rate of 256 Hz. The research included 122 men, both healthy and drunk. The research, which were open to all participants, used visual stimuli from the Snodgrass and Vander Wart collections. [76] Two stimulants were employed, one lasting 300 milliseconds and the other lasting 1.6 seconds. If S1 and S2 are the same, they were actually asked to answer the question. Certain scenarios demanded the use of only one trigger [77].

2.13.6 Bonn Database

The Bonn EEG dataset was collected by Andrzejak et al. [78] at the University of Bonn in Germany. This dataset consists of five sets of EEG signals, A, B, C, D, and E. Each set has 100 single-channel EEG fragments of 23.6 s, sampled at a rate of 173.61 Hz, giving each set a total of 4097 samples. All sets have a spectral bandwidth between 0.5 and 85 Hz. The identical 128-channel amplifier equipment was used for all of the recordings. For Sets A and B, information was collected from five healthy individuals using standardized surface EEG recordings while their eyes were open and closed was collected from five healthy individuals using standardized surface EEG recordings while their eyes were open and closed, respectively, for Sets A and B. Five epileptic individuals with a pre-surgical epilepsy diagnosis provided Sets C, D, and E. Set C and set D were recorded during epilepsy free interval and set E was recorded during the occurrence of epileptic seizures [79].

Table 2.1: Overview of Bonn Dataset.

Set	Patients	Setup	Phase
A(O)	Healthy	surface EEG	open eyes
B(Z)	Healthy	surface EEG	closed eyes
C(F)	Epilepsy	intracranial EEG	Interictal
D(N)	Epilepsy	intracranial EEG	Interictal
E(E)	Epilepsy	intracranial EEG	Seizure

In the current thesis, Bonn dataset in the proposed methods is used.

2.14 Performance Evaluation

This section provides some important performance measures for assessing the suggested strategy. They are characterized as follows:

I. **Compression Ratio (CR):** this metric is defined as follows:

$$\textbf{CR(\%)} = \left(1 - \frac{E_{Compr}}{E_{Or}}\right) *100, \ldots\ldots\ldots \quad \textbf{(2.1)}$$

Where E_{Compr} denotes the size of the compressed EEG data after using the advised method, and E_{Or} denotes the size of the uncompressed EEG data.

II. **Compression and Decompression Processing Time (Seconds):** This is the overall amount of time required for the compression and decompression procedures.

III. **Size of Sent Data(KB):** The compressed EEG data supplied in KB form from the fog gateway to the cloud server makes up the amount of transmitted data.

IV. **Compression Power (CP):** typically known as the data compression ratio, is defined as the proportion of the uncompressed size (E_{Or}) to the compressed size (E_{Compr}) of data. The following is a definition of CP:

$$\textbf{CP} = \frac{E_{Or}}{E_{Compr}}, \ldots\ldots\ldots\ldots \textbf{(2.2)}$$

V. **Average Compression Power (ACP):** refers to the average of CP of all the dataset records (Z, F, N, O, S). It can be define as follows:

$$\textbf{ACP} = \frac{CP^{Z}+CP^{F}+CP^{N^{+}}+CP^{O}+CP^{S}}{5}, \ldots\ldots\ldots\ldots \textbf{(2.3)}$$

The parameters CP^Z , CP^F , CP^N , CP^O, and CP^S refer to the compression power of dataset records Z, F, N, O, and S respectively.

Summary

This chapter covers the fundamentals of the WBSN architecture, such as communication levels, wireless body sensor networks medical applications, technologies, hardware devices types and characteristics and challenges. The definition of fog computing and its characteristics with benefits are investigated. The EEG biosensors will be used as an example of a WBSN. The notion of data clustering with its types and data compression, as well as its two primary forms, lossless and lossy are demonstrated. The EEG Datasets types are explained. Finally, the performance evaluation metrics are illustrated.

Chapter Three

The Proposed Approaches

3.1 Introduction

The possibility of transferring patient health records through the network has been expanded. This has resulted in high network traffic and latency, especially considering that the majority of transferred data are health records, which account for a significant portion of data volumes. Health EEG data compression techniques are presented to decrease the massive amount of network transmitted data. This chapter presents three efficient lossless data compression models for lowering network traffic volume and enhancing the performance of the IoMT networks. The computational operations of proposed systems are supposed to be executed in the fog gateway to reduce IoMT traffic and improve IoMT application response times. Lastly, transmit compressed health data to the cloud for medical health.

3.2 EEG Compression Approaches

Three lossless EEG data compression techniques are suggested in this part to reduce the amount of EEG data sent to the cloud and improve network performance. The suggested methods are meant to be put into practice at the fog gateway, which is meant to be situated close to the EEG data generators.

3.2.1 The LEDaC Technique

This section introduces the Lossless EEG Data Compression (LEDaC) for Fog computing based IoMT networks. LEDaC reduce the volume of EEG data by compressing them before sending them to the cloud data center. This will improve the performance of the IoMT networks. The proposed LEDaC method is shown in Figure (3.1).

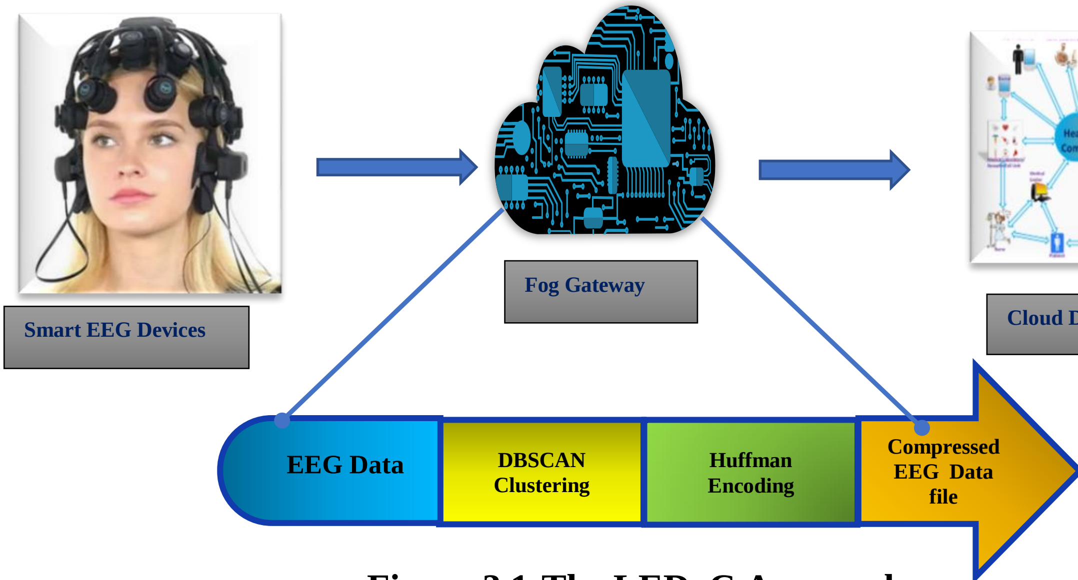


Figure 3.1:The LEDaC Approach

The EEG electrodes are a group of sensors that are placed on the patient's head using EEG headset to record the activity of the brain and wirelessly transmit it to the fog gateway. The fog gateway processes and compresses the collected EEG data before sending it to a cloud platform for additional analysis. The EEG data are gathered periodically from the headset biosensors at the fog gateway. Algorithm (3.1) refers to the LEDaC technique.

Algorithm (3.1) will be executed on a periodic basis in the fog gateway. During each period, the suggested LEDaC technique would compress the accumulated EEG data at the Fog gateway and then forwarded it to the cloud. After acquiring EEG data from the Fog gateway, the suggested LEDaC method transforms the EEG data into several groups based on their similarity using DBSCAN clustering.

Algorithm 3.1: LEDaC algorithm

Input: *$D= d_1, \ldots, d_n$, n, mS: the fewest points required to create a dense region, eps: the range of neighborhood*

Output: *EF : encoded file*

```
CR ← Φ;
for k ← 1 to n do
    if ( d_k is visited ) then
        Keep on going to the next EEG value d_{k+1};
    end
    else
        Confirm as visited the EEG data value d_k;
        nS ← eps-neighborhood EEG values of d_k;
        if ( Sizeof(nS) < mS) then
            Highlight the value of the EEG as Noise;
        end
        else
            CR ← NC;  // NC is a new cluster
            Call Expansion of Cluster Function (d_k, nS, CR, mS);
        end
    end
end
{G_1, G_2, . . . , G_β } ← AllocateGroups (CR, β); // β is total number of clusters
FL ← " " // empty file;
for i ← 1 to β do
    GI ← BringIndexes(G_j);
    FL ← G_1 ∪ GI;
end
EF ← Huffman Compression Algorithm (FL) ;
return EF;
```

Algorithm 3.2: Expansion of Cluster

Input: *d, nS, CR, mS, and eps*
Output: *CR: current cluster*

$CR \leftarrow CR \cup d$;
for $j \leftarrow 1$ *to* nS *do*
    if (d_j^* *is not visited*) *then*
        Confirm as visited the EEG value d_j^*;
            $nS^* \leftarrow$ *eps-neighborhood EEG values of* d_j^*;
            if (*Sizeof*($nS^* \geq mS$) *then*
                $nS \leftarrow nS + nS^*$;
            end
    end
    if (d_j^* *does not found at any cluster*) *then*
    $CR \leftarrow CR \cup d_j^*$;
    end
end
return CR;

These EEG data are represented as a series D= d1, ..., dn, where n is the total number of EEG data during one period (8194 EEG data values). The EEG data series D is clustered using DBSCAN clustering technique into several clusters of similar or identical values. Clusters can be identified using the density of points. DBSCAN grows clusters using a density-based connectivity analysis. The cluster is described as the most densely connected points. The fundamental idea behind density based clustering is that the neighborhood of a specified radius (eps) must involve considerable number of objects (MinPts), i.e., the cardinality of neighborhood must exceed up a particular threshold. The DBSCAN technique was designed to effectively locate clusters and noise in a dataset.

DBSCAN uses two parameters: (eps) and the lower number of points necessary to produce a cluster (Minpts). It starts at a random location that has never been visited before. This point's -neighborhood is retrieved, and if it contains enough points, a cluster is formed. Otherwise, the point is classified as noise. The DBSCAN algorithm excels at dealing with large datasets, noise, and classifying clusters of different sizes and shapes. For example, when it comes to complex object clustering, density-based clustering has proven to be especially effective for assessing a huge volume of heterogeneous, complicated data.

Steps (1-17) demonstrate the EEG data into groups of similar or identical data using the DBSCAN method. The expanding function (Algorithm 3.2) is included in Algorithm(3.1). The function AllocateGroups() in step 18 returns the clustered groups of the EEG values and keeps them in lists such as G1, G2, . . . , Gβ. The number of groups is generated automatically based on the value of the parameter eps in the DBSCAN algorithm. Steps 20-23 return the indexes of the EEG data in each group and one representative EEG value and saved it into a file. This file will be passed to the Huffman encoding to encode its values (see step 24) using Algorithm (2.1) in Chapter 2. The Huffman encoding algorithm is a powerful compression algorithm that can be used to encode the data efficiently to decrease its volume without losing any information. Each encoded file will be transmitted to the cloud for additional processing and analysis. The cloud data center receives the encoded files and decodes them using Algorithm (2.2) in Chapter 2. Each decoded file will print the representative EEG value in each index position in the EEG time series to reconstruct the original EEG data.

3.2.2 The ECoT Technique

The Effcient Compression Technique (ECoT) for reducing the transmitted EEG data without loss in fog computing based IoMT networks is introduced in this section. Before transmitting EEG data to the cloud platform, ECoT compresses it to decrease its volume so as to improve the performance of IoMT networks. The ECoT technique is illustrated in Figure (3.2) .

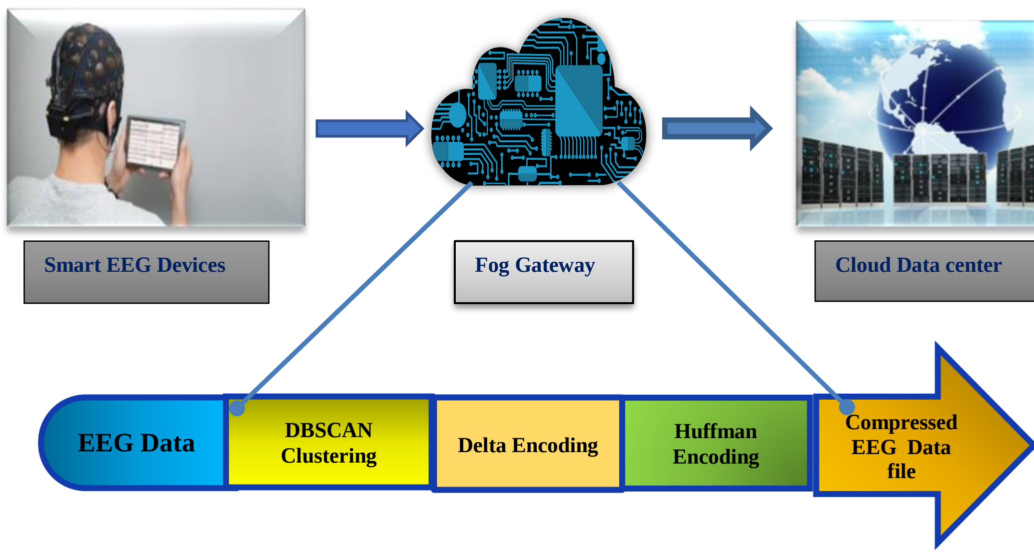


Figure 3.2: The ECoT Approach

Before sending the gathered EEG data to a cloud platform for additional analysis, each period, the fog gateway processes and compresses two records of EEG signals, which can be registered as X = x1, x1, ..., xβ, where β is the sum of EEG data for two records, which equals 8194 EEG values. The ECoT technique is presented in Algorithm (3.3).

Algorithm 3.3: ECoT Technique

***Input**: X= x1,. . . ,xβ, β, minSamplings: the smallest number of points needed to establish a dense area, eps: the neighborhood range*

***Output**: F^{DH}: compressed file*

Cluster ← Φ;
for p ← 1 to β do
    if (X_p is visited) then
        Continue to next EEG data value X_{p+1};
    end
    else
        Indicate EEG data value X_p as visited;
        nbrSamplings ← EEG values in the eps-neighborhood of X_p ;
        if (Size(nbrSamplings) < minSamplings) then
            indicate EEG data value as Noise;
        end
        else
            Cluster ← NovelCluster ;
            Call Cluster Expansion Algorithm (X_p, nbrSamplings, Cluster, minSamplings);
        end
    end
end
{C 1 , C2 , . . . , CK} ← *AllocateGroups* (Cluster, K);
F^{DH} ← K;
for j ← 1 to K do
    INDX ← AssignIndices (C^j);
    Delta ← Delta Encoding Algorithm(INDX);
    F_{DE} ← C1 ∪ Delta;
    F^{compr} ←Huffman Compression Algorithm(F_{DE}) ;
    $F^{DH} \leftarrow F^{DH} \cup F^{compr}$;
end
SendToCloud(F^{DH}) ;
return (F^{DH});

In the fog gateway, Algorithm (3.3) will be implemented periodically to compress and transmit the compressed file to the cloud. The ECoT technique uses DBSCAN clustering to divide the data into several groups based on their similarity.

Steps 1-17 illustrate the DBSCAN strategy is demonstrated with its expanding function (Algorithm 3.2) in Algorithm (3.3).the DBSCAN strategy will be as mentioned above in the previous method. Steps 20-26 are responsible for compressing each cluster and sending it to the cloud. The function AssignIndices() takes the EEG data of each cluster and returns their indices to save them in the vector INDX (see step 21).

In step 22, the Delta encoding algorithm is applied to the vector of indices INDX and saves the differences in the vector EINDX. The Delta encoding is applied to the indices of the data of each cluster. The result is a vector of differences between indices. One representative EEG value of the cluster is inserted at the beginning of the vector EINDX. The main objective of saving the indices values as differences (deltas) between sequential indices values is to reduce the range (variance) of values when neighbour data are correlated, allowing a lower bit usage for the same data. This resulted vector will be passed to the Huffman encoding. In step 24 of Algorithm (3.3),

The Huffman compression Algorithm (2.1) in Chapter 2 is employed to compress the resulting file by the Delta encoding technique to decrease the quantity of data uploaded to the cloud**.** A leaf node is constructed for each EEG value and positioned into the priority queue. In Algorithm (2.1), step 6 constitutes a loop that would operate as long as the queue has more than one node. Steps 7-11 will eliminate the nodes with the highest priority from the queue. It combines the frequencies of these two nodes to produce a new node that is the child of these two nodes. After that, a queue will receive the new node. The root node of the tree is dormant and the tree is complete. In step 25 of Algorithm (3.3) the compressed data of each cluster will added to the final file. Finally, the compressed data of the period will be sent to the cloud.

In the other side of the network (cloud), the received compressed file will be decompressed to get the original EEG data. First, the compressed file produce k compressed vectors that represents the encoded indices of each cluster. Each encoded indices vector is decoded by Huffman Decompression Algorithm (See Algorithm 2.2) in Chapter 2. The result is a vectors of differences with their EEG representative values.

The vector of differences between the indices of EEG data of each cluster is decoded using Delta decoding algorithm to get the original indices of the EEG values for each cluster. Finally, the data of each group will be reconstructed by putting the EEG value in every index inside the cluster.

3.2.3 The NoLEDaC Technique

The suggested Novel Lossless EEG Data Compression (NoLEDaC) approach enabled by fog computing in the Internet of Health Things is presented. The NoLEDaC technique compresses EEG data to reduce the amount of data transmitted to cloud data centers and improve the IoT network's performance.The NoLEDaC technique offered by fog computing is shown in Figure(3.3). The EEG biosensors are implanted on the patient's scalp to capture the activity of brain and send it wirelessly to the fog smart gateway. Patients' data is collected via the headset of the patient and then delivered to the gateway of fog layer regularly. The acquired EEG data is processed and compressed by the proposed NoLEDaC approach at the fog gateway. Then, it is transferred for further analysis to a distant cloud (See Figure 3.3). To begin with, the EEG data is gathered from the EEG electrodes in the headset of the patient and received by the fog gateway periodically.

Each period, the fog gateway processes two recordings of EEG data readings, which can be written as Y = y_1, y_1, ..., y_β, where β is the sum of the two recordings' EEG data readings, which equals 8194 EEG values. Algorithm (3.4) refers to the proposed NoLEDaC approach.

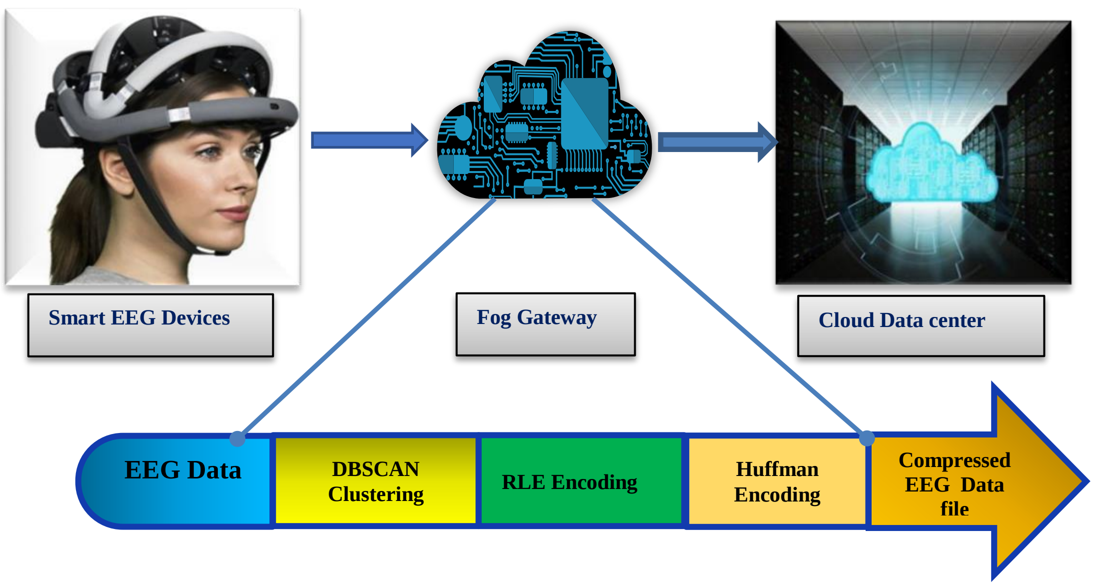


Figure 3.3: The NoLEDaC Approach

Algorithm (3.4) will be implemented in the fog gateway periodically. In each period, The EEG data collected at the Fog gateway would be compressed and transferred to the cloud using the suggested NoLEDaC method. After the Fog gateway has received the EEG data, the proposed NoLEDaC approach uses DBSCAN clustering to divide the EEG data into several groups according to their similarity. Steps 1-17 show the DBSCAN technique is illustrated with its expanding function (Algorithm 3.2) in Algorithm (3.4).

Algorithm 3.4:NoLEDaC Technique

Input: *Y= y1,. . . ,yβ, β, minSamplings: the smallest number of points needed to establish a dense area, eps: the neighborhood range*

Output: F^{compr}**:** *compressed file*

Cluster ← Φ;
for p ← 1 to β do
if (y_p is visited) then
Continue to next EEG data value y_{p+1};
end
else
Indicate EEG data value y_p as visited;
nbrSamplings ← EEG values in the eps-neighborhood of y_p ;
if (Size(nbrSamplings) < minSamplings) then
indicate EEG data value as Noise;
end
else
Cluster ← NovelCluster ;
Call Cluster Expansion Algorithm (y_p, nbrSamplings, Cluster, minSamplings);
end
end
end
$\{C^1, C^2, \ldots, C^K\}$ ← *AllocateGroups* $(Cluster, K)$;
F_{RLE}← " " // *empty file;*
for $j \leftarrow 1$ **to** K **do**
EEGvalue, Frequency ← RLE Algorithm(C^j);
$F_{RLE} \leftarrow F_{RLE} \cup$ *EEGvalue, Frequency;*
end
$F^{compr} \leftarrow$ *Huffman Compression Algorithm;*
return F^{compr};

The approach starts by selecting one EEG data value y_p and thereafter calculating the distance between that value and the remainder of the EEG data values in the dataset. If the distance between the EEG data value y_p and any other EEG data value in the dataset is less than or equal to *eps*, they are considered neighbours.

A new cluster is established if the number of EEG data values in the proximity of EEG data value y_p is more than or equal to $min_{Samplings}$; otherwise, these EEG data values are marked as noise. This means that noisy EEG data values in the neighborhood range of the newly selected EEG data value y_p can afterwards be included in other clusters if they match the $min_{Samplings}$ criterion. The DBSCAN technique will then determine if it is possible to extend this cluster or whether it should select an EEG data value from outside the present cluster.

The checking is done by ensuring that both the $min_{Samplings}$ and distance criteria are met for each EEG data value in the cluster’s range. If these requirements are met, the DBSCAN technique expands this cluster to include all EEG data values in the range of neighborhood of y_p. If the cluster extends to the needed $min_{Samplings}$, the cluster will be stopped and each EEG data value will be labelled as a visited EEG data value.

The DBSCAN algorithm then selects another nonvisited EEG data value from the dataset and continues the process. If there are no EEG data marked as "not visited," the DBSCAN technique will be terminated. In step 18, the function *AllocateGroups* () returns the cluster number with the EEG data values for each cluster and then saved in sets like $C^1, C^2, \ldots, C^K$. The DBSCAN algorithm automatically generates the number of clusters based on the *eps* parameter. These cluster sets are transmitted to the RLE algorithm. Then, the RLE encoding method is applied to the EEG data of each cluster to provide a file includ- ing the EEG values with their frequencies. As a result, after compressing the data by RLE, it reduces the number of EEG data while maintaining data accuracy.This file is then compressed using Huffman encoding. The final compressed file will be transmitted to the cloud platform (See steps 20-23). In Algorithm (3.4),

The steps 1-17 require O(β^2) of time complexity, where β is the size of the collected data per period. The time requirements are lowered to O($\beta log\beta$) when spatial indexing is applied. In step 24 of Algorithm (3.4), the Huffman compression algorithm (See algorithm 2.1 in Chapter 2) is employed to by the proposed NoLEDaC approach to compress the resulting file by the RLE technique to decrease the quantity of data trans- ferred to the cloud. A leaf node is created for each EEG value and placed into the queue of priority. Step 6 comprises a loop that would run as long as the queue has more than one node. Steps 7–11 will remove the nodes with the highest priority from the queue. It combines the frequencies of these two nodes to produce a new node that is the child of these two nodes. After that, a queue will receive the new node. The root node of the tree is dormant and the tree is complete. Algorithm (3.4) has a time complexity of O($\beta log\beta$). Hence, the overall time complexity of the proposed NoLEDaC approach takes O($\beta log\beta$).

The prefix code sequence is exchanged for a specified byte value in the decompression method. It is accomplished by tracking the tree as each bit from the input series is collected. When the leaf node is arrived by the traversing, the byte value is finished. The desired EEG value is expressed by the leaf value.

Summary

In this chapter,Three lossless EEG data compression approaches are proposed that are supposed to be implemented at the fog gateway to reduce the redundant data after compressing them at the fog gateway and then send them to the cloud for further analysis and archiving. First, the Lossless EEG Data Compression (LEDaC) is introduced for Fog computing based IoMT networks. Then, an Effcient Compression Technique (ECoT) for reducing the transmitted EEG data without loss in fog computing based IoMT networks is introduced. finally, a Novel Lossless EEG Data Compression (NoLEDaC) approach enabled by fog computing in the Internet of Medical Things is presented. The main objective of these proposed approached is to reduce the volume of EEG data by compressing them at the fog gateway before sending them to the cloud data center. This is contributed in improving the performance of the IoMT networks.

Chapter Four

Performance Evaluation and Results

4.1 Introduction

In this chapter, the results of testing the proposed compression techniques: the LEDaC, ECoT, and the NoLEDaC techniques are presented.

Multiple experiments based on fundamental metrics and real dataset are used to evaluate the performance of the proposed compression strategies. A comparison with other previously published compression algorithms has been conducted to evaluate the effectiveness of the suggested compression techniques. It has been found that the suggested methods have completely surpassed them.

The computation operations of the proposed approaches have been done on (DELL) laptop computer that have the following characteristics:

- **Processor** : 11th Gen Intel(R) Core(TM) i7-1165G7 @ 2.80GHz 2.80 GHz .
- **Memory** : 8.00 GB.
- **System type** : 64-bit operating system, x64-based processor.

and a python programming language version (3.9) was employed. The proposed approaches are tested by using public Bonn University dataset which mentioned in Chapter Two, section (2.13.6).

4.2 The LEDaC Technique Results

The primary goal of the proposed LEDaC solution is to reduce communication by reducing the total amount of data transferred to the cloud. This is performed by compressing the recorded EEG data before delivering it to the cloud via the fog layer's gateway. The proposed LEDaC's performance is evaluated utilizing a range of assessment indicators, including compression ratio, computational time for compression and decompression, and the sent data to the cloud.

The eps and mS values were set at 0.20 and 5, respectively. The proposed LEDaC method is compared to current approaches such as JPEG2000, AC, 1-D SHORTEN, and 2-D SPIHT + AC [22], as well as 2-D SPIHT and 1-D SPIHT [21]. The next subsections present a series of tests based on the assessment criteria employed by the proposed LEDaC approach. The Min, Avg, and Max values in these experiments relate to the minimum, average, and maximum values of computed performance assessment metrics for 50 periods, where these metrics are generated for each period and then the Min, Avg, and Max values for them are considered based on 50 times.

4.2.1 The Volume of Sent EEG Data

These experiments demonstrate the effect of the proposed LEDaC method on the size of EEG data delivered to the cloud. Figure (4.1) depicts the communicated data in KB to the cloud platform after implementing LEDaC for various dataset records (Z, N, S, F, O). Figure (4.1) shows the transmitted data in KB.

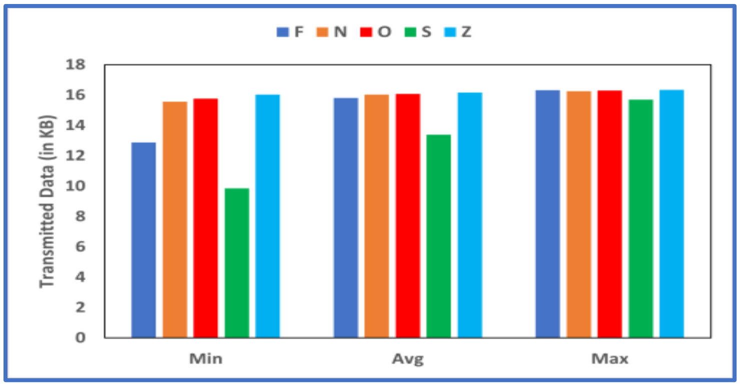
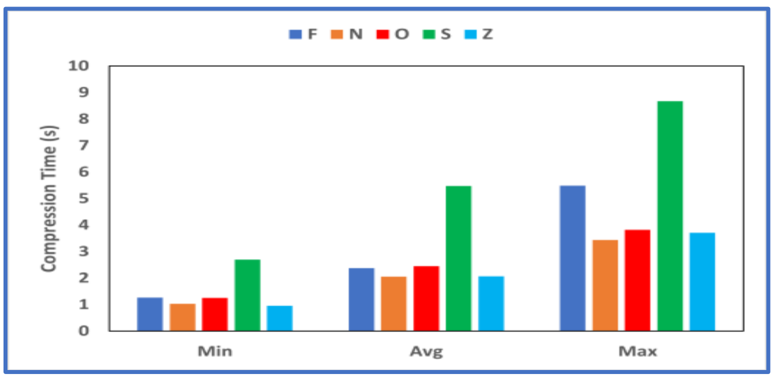


Figure (4.1) :The Transmitted EEG Data in KB(LEDaC)

4.2.2 The Compression Time

The compression time required to compress the EEG data per period is assessed in this study. Figure (4.2) shows the time of compression for different dataset records (Z, N, S, F, O).



Figure(4.2): The Compression Time(LEDaC)

It can be observed from the Figure (4.2) that the proposed LEDaC technique compress the EEG data of each period with a suitable time, where it requires from 0.95 up to 8.67 seconds to compress the data of one period for different dataset records (Z, N, S, F, O).

4.2.3 The Decompression Time

In this research study, we assess the proposed LEDaC technique from the decompression time point of view. Figure (4.3) shows the time of decompression for different dataset records (Z, N, S, F, O).

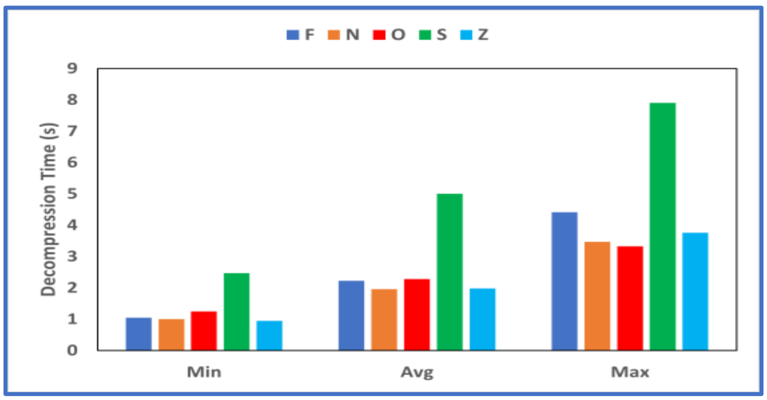


Figure (4.3): The Decompression Time(LEDaC)

It can be noticed from the Figure (4.3) that the proposed LEDaC technique decompresses the received compressed file at the fog gateway with a suitable time, where it needs from 0.95 up to 7.9 seconds to decompress the received compressed file of one period for different dataset records (Z, N, S, F, O).

4.2.4 The Compression Ratio

In order to evaluate the suggested LEDaC technology, the compression ratio is computed in this experiment. Figure (4.4) shows the compression ratio for different dataset records (Z, N, S, F, O).

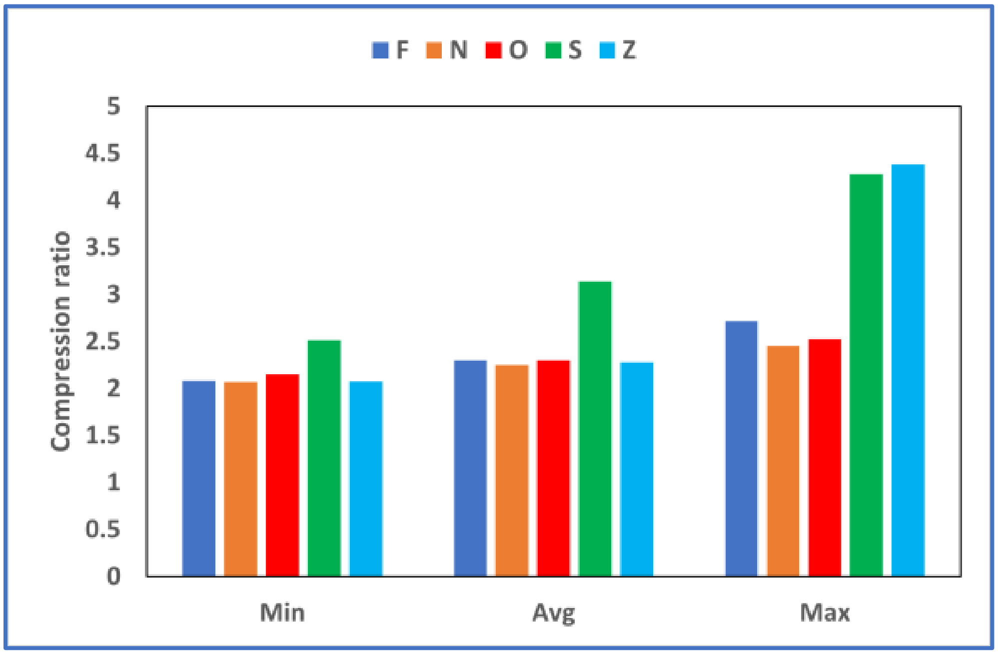


Figure (4.4): The Compression Ratio(LEDaC)

It can be seen from the Figure (4.4) that the proposed LEDaC technique introduce a good compression ratio, where it compresses the EEG data per period from 2.1 up to 4.39 for different dataset records (Z, N, S, F, O).

The proposed LEDaC technique is compared with various methods (as shown in Table 4.1) from the compression Point of view. It can be seen from the results in Table (4.1) that the proposed LEDaC technique It outperforms other methods by offering better compression ratio for different dataset records (Z, N, S, F, O).

Table (4.1): The Comparison of Compression Ratio for Different Methods.

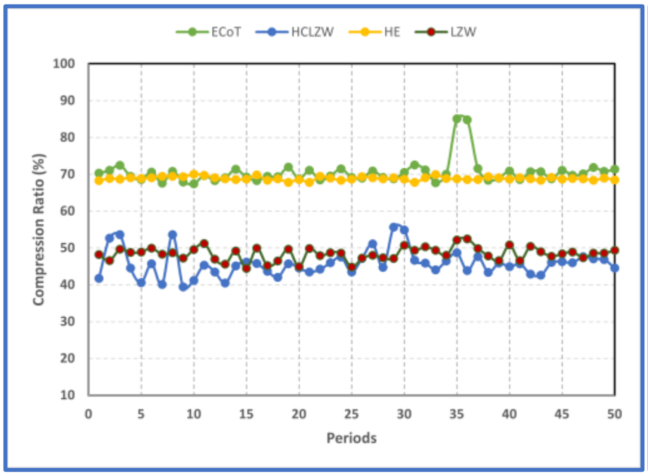
Dataset	LEDaC	2-D SPIHT + AC	JPEG2000	2-D SPIHT	1-D SHORTEN	Ac	1-D SPIHT
F	2.30	2.19	2.13	2.18	1.34	1.51	2.02
N	2.25	2.23	2.15	2.15	1.27	1.55	1.99
O	2.30	1.84	1.87	1.86	1.24	1.46	1.77
S	3.14	1.44	1.48	1.44	1.27	1.15	1.42
Z	2.28	2.01	1.9	1.99	1.18	1.16	1.85
Average	**2.454**	**1.942**	**1.906**	**1.924**	**1.26**	**1.366**	**1.81**

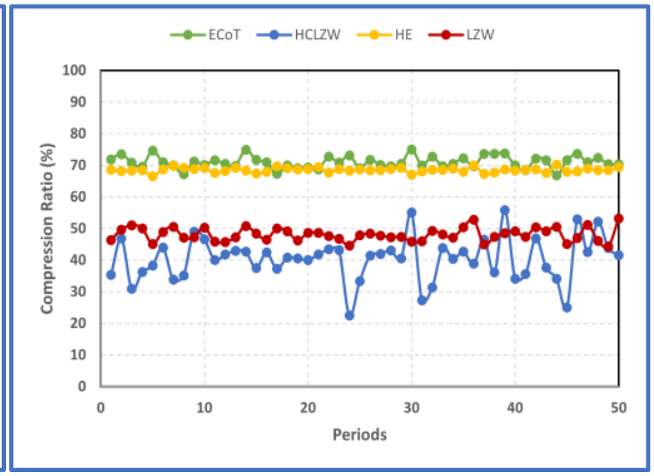
4.3 The ECoT Teqnique Results

The proposed ECoT technique's performance is evaluated using several metrics such as compression ratio, compression power, compression and decompression computational time, and data transported to the cloud. The suggested ECoT technology is contrasted with existing technologies such as HCLZW lossless compression combines hierarchical clustering with LZW lossless compression. The LZW approach is the original LZW method, whereas the HE is the Huffman Encoding [33].

4.3.1 The Compression Ratio

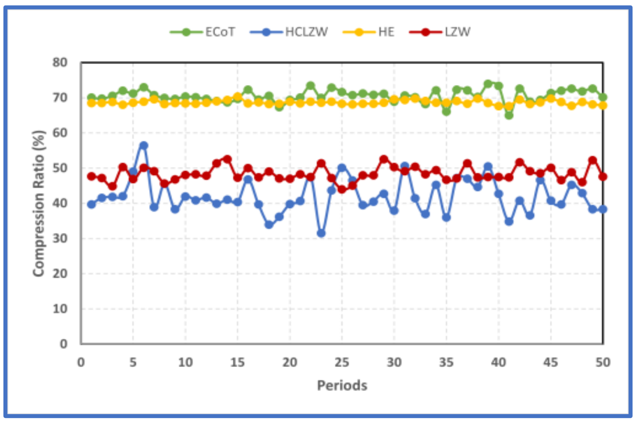
It is an important factor that impacts the network performance**.** The ECoT approach reduces the EEG data by 65 to 85 percent for different records. Comparatively, the other methods reduce the EEG data's size by 40 to 54 percent, 65 to 70 percent, and 14 to 56 percent for LZW, HE, and HCLZW, respectively, for all recordings. The findings demonstrate that by offering a greater compression ratio, the suggested ECoT outperforms previous approaches in reducing the transmitted data to the cloud. Figure (4.5) illustrates the compression ratio for different records like S, N, Z, F, and O

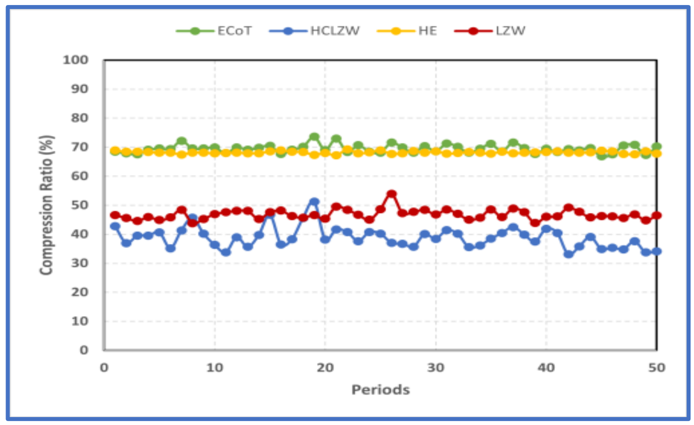




(a)

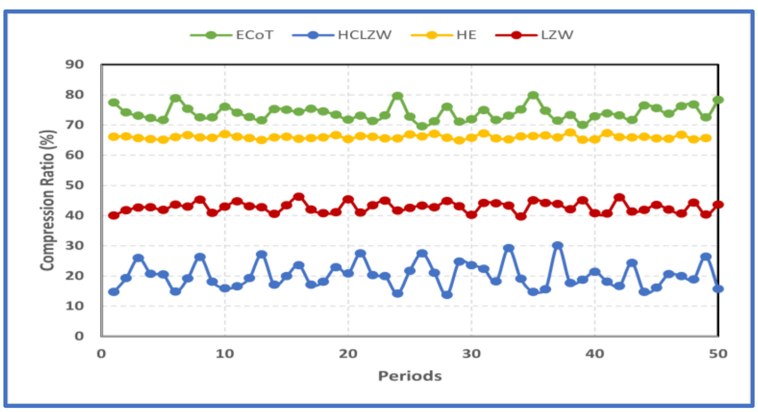
(b)





(c)

(d)



(e)

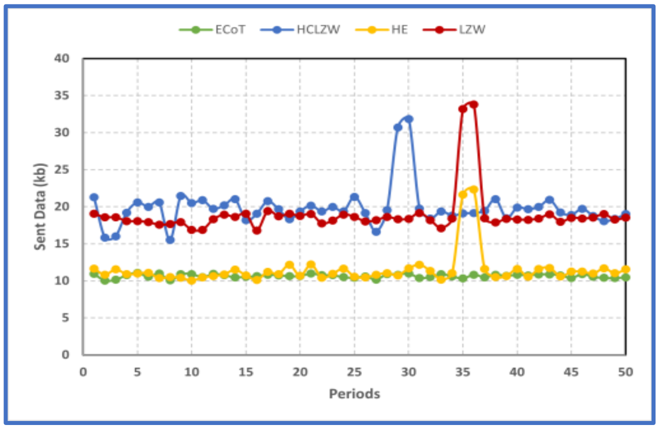
Figure 4.5: The Compression Ratio(ECoT). (a) Z, (b) F, (c) N, (d) O, (e) S.

4.3.2 The Sent EEG Data

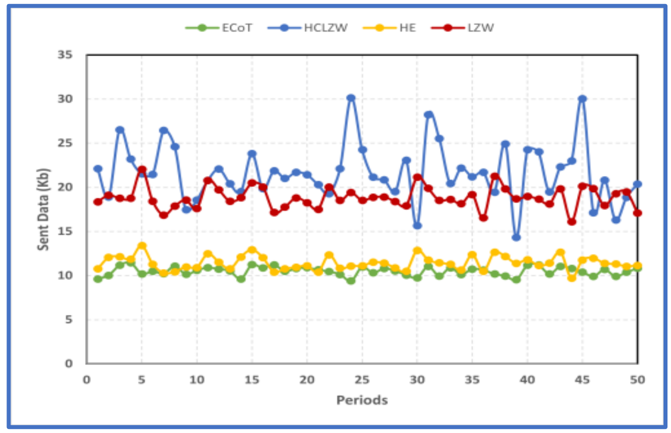
This experiment looks at how the suggested ECoT technique affects EEG data delivered to a cloud server. The sent data for a range of procedures utilizing various records is depicted in Figure (4.6). The results show that, when compared to existing approaches, the suggested ECoT technique can give EEG data in a more compact manner. The recommended ECoT decreases the amount of delivered EEG data (in KB) from 9.4 to 12.0 for varied recordings, according to the results. While the other techniques reduced the volume of transmitted EEG data (in KB) for various recordings from 16 to 34, 10 to 22.3, and 14 percent to 37 percent for LZW, HE, and HCLZW, respectively. As a result, ECoT decreased transmitted EEG data more efficiently than other methods.

4.3.3 The Compression Time

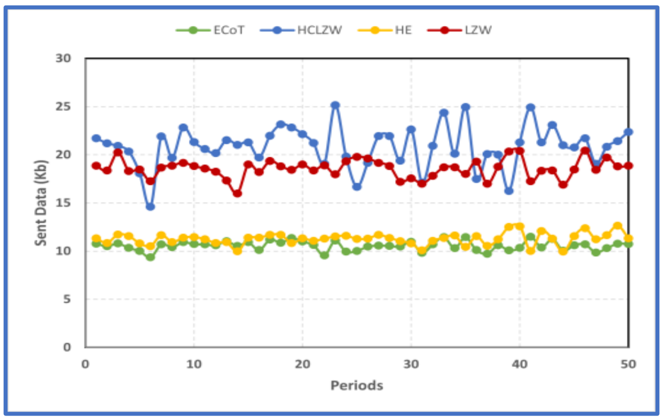
This experiment examines the effect of the proposed ECoT on EEG data compression time. Figure (4.7) depicts the compression time for several methods employing distinct records N, O, Z, F, and S. In comparison to HCLZW, HE, and LZW, the ECoT provides less compression time per period for different types of data, according to the results. It requires somewhat more time to compress than HE in F and S records. The proposed ECoT enhances network performance by decreasing the amount of EEG data transmitted and increasing the compression ratio, both of which have a significant impact on network performance.



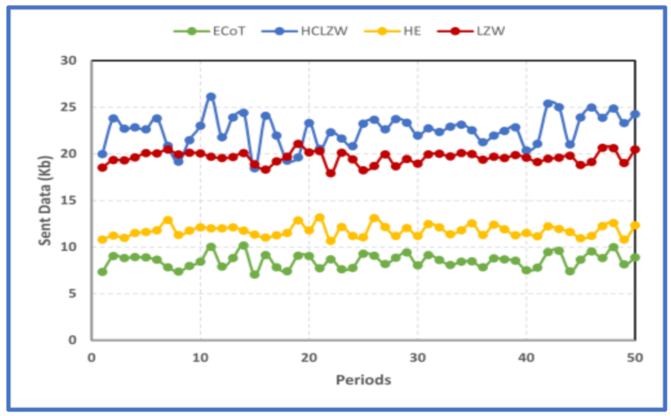
(a)



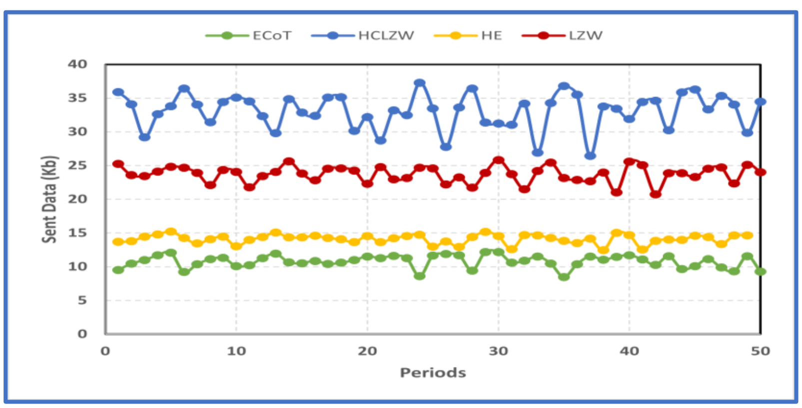
(b)



(c)

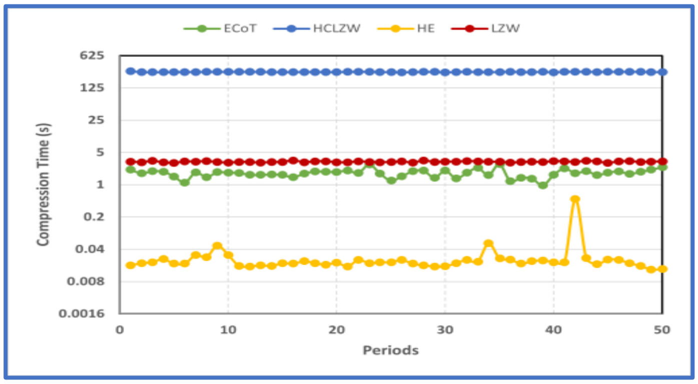


(d)



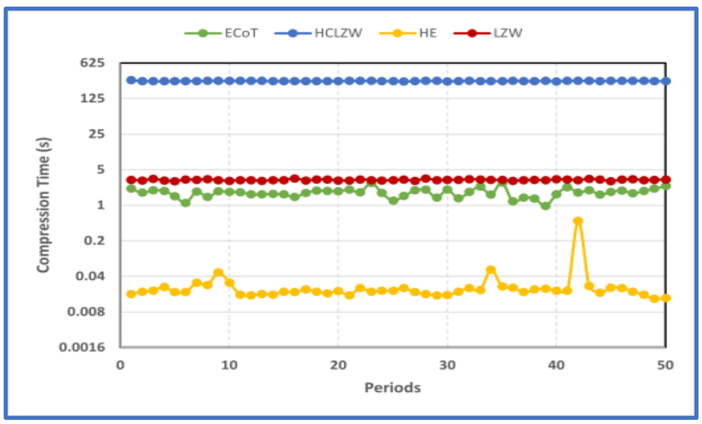
(e)

Figure 4.6: The Sent EEG data(ECoT). (a) Z, (b) F, (c) N, (d) O, (e) S.

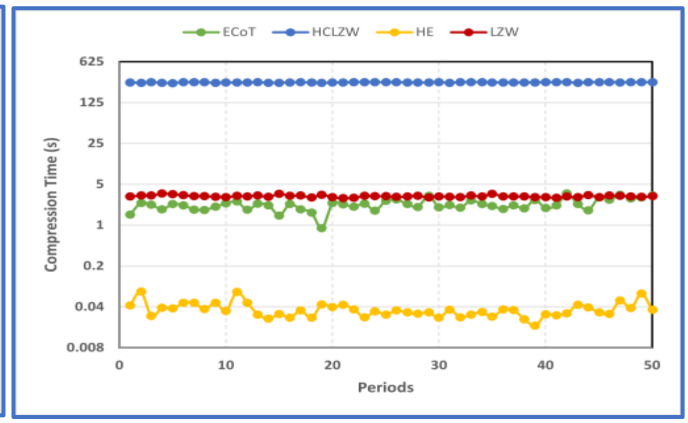


(a)

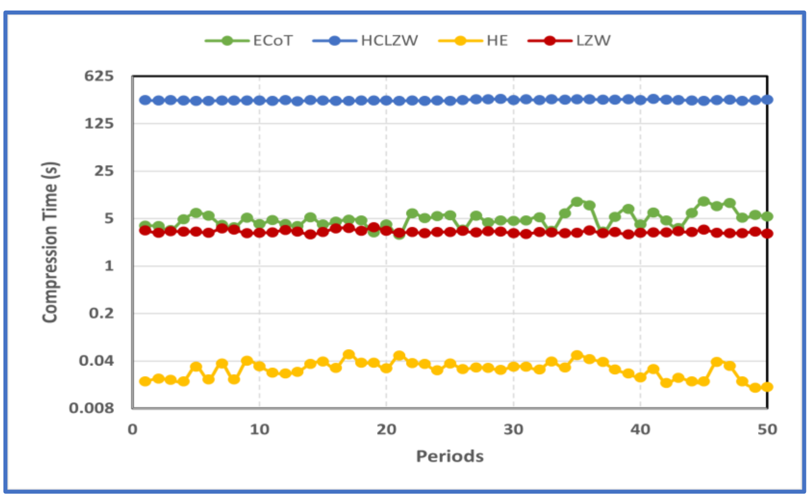
(b)



(c)



(d)

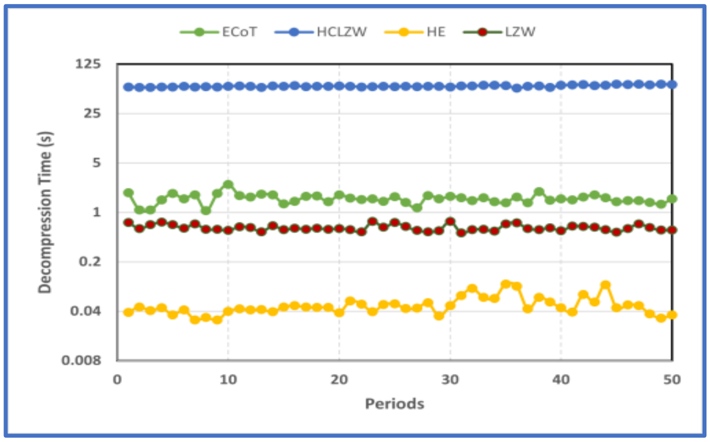


(e)

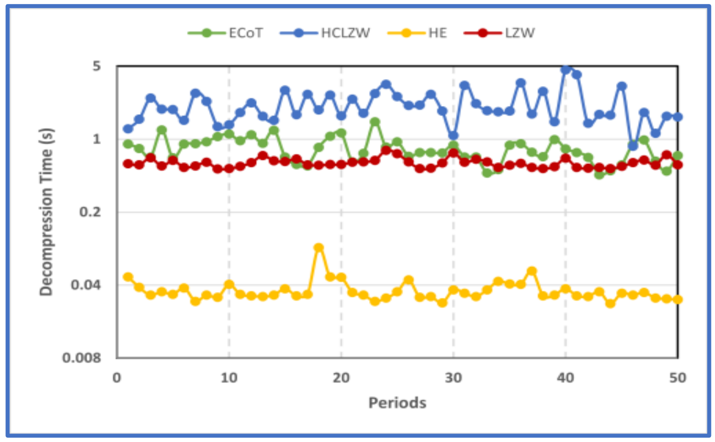
Figure 4.7: The Compression Time(ECoT). (a) Z, (b) F, (c) N, (d) O, (e) S.

4.4.4 The Decompression Time

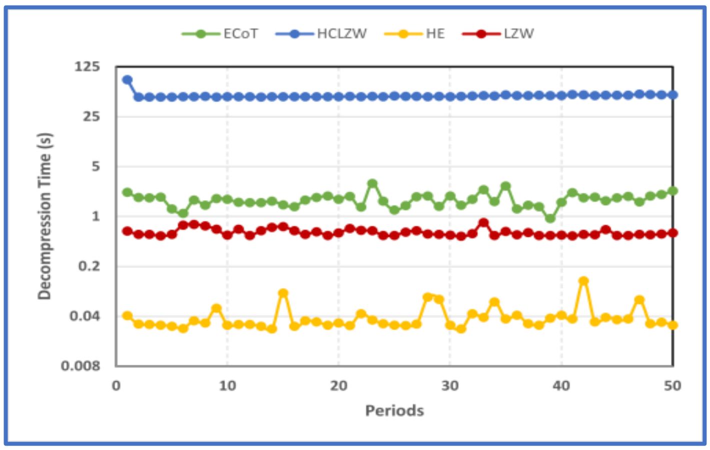
The impact of the proposed ECoT on the decompression time of ”EEG data” is studied in this section. The decompression time for many approaches utilizing different records like N, O, Z, F, and S is illustrated in Figure(4.8). In comparison to HCLZW, HE, and LZW, the results demonstrate that the ECoT technique gives lower decompression time in the case of record N, while it introduces silghter higher decompression time than HE, and LZW in the records S, F, O, and Z.



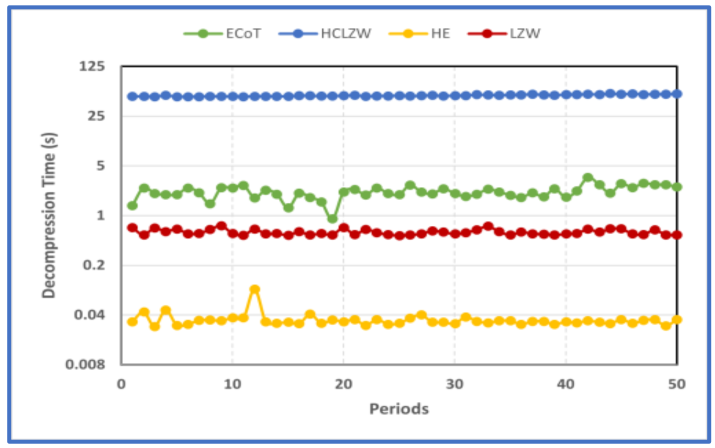
(a)



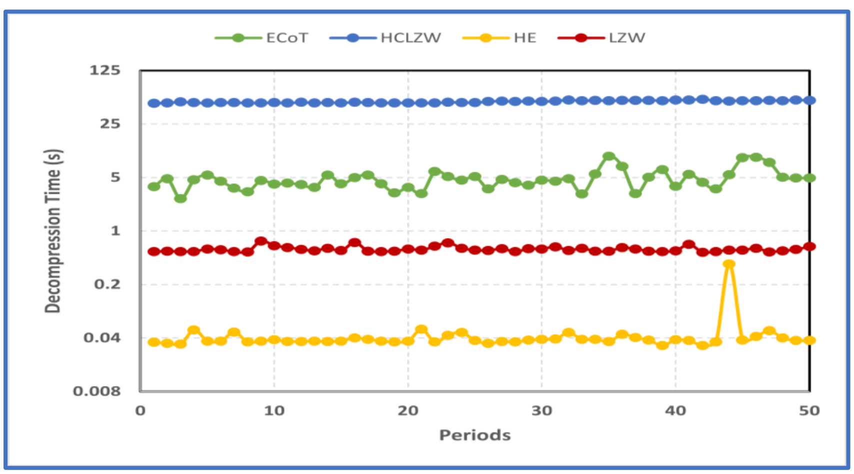
(b)



(c)



(d)



(e)

Figure 4.8: The Decompression Time(ECoT). (a) Z, (b) F, (c) N, (d) O, (e) S.

4.4.5 Compression power and Average compression power

The proposed ECoT is evaluated using both the compression power and the average compression power performance metrics, which are defined in Chapter Two in paragraph(2.13), and compared to some existing techniques in the related works like 1-D SPIHT [21], 2-D SPIHT + AC [22], 1-D SHORTEN [22], JPEG2000 [22], 2-D SPIHT [21], and AC [22].

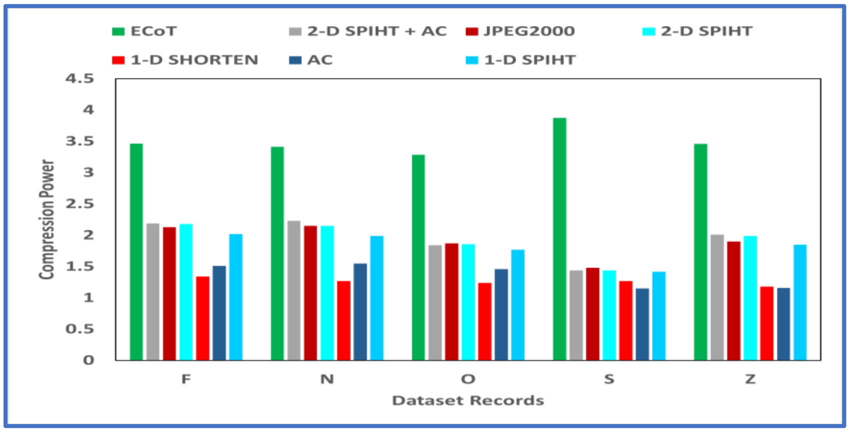


Figure 4.9: The Compression Power(ECoT)

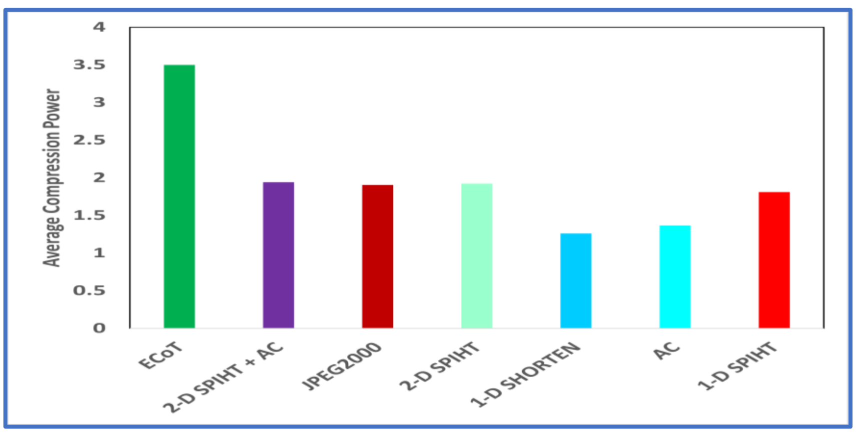


Figure 4.10: The Average Compression power(ECoT)

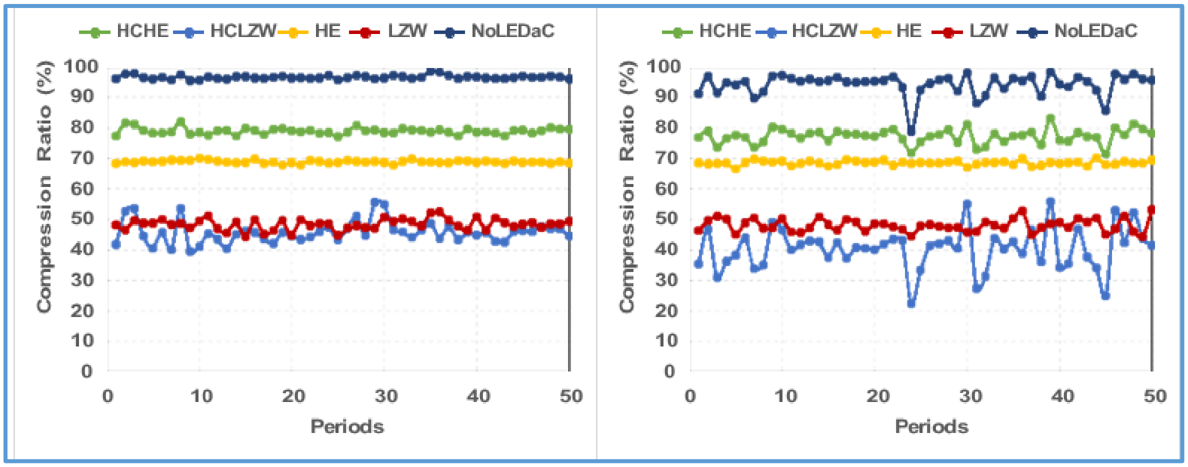
Figures (4.9) and (4.10) show that the ECoT technique outperform the alternative methods in terms of compression power, resulting in a reduction in the sent EEG data and an improvement in network performance.

4.5 The NoLEDaC Technique Results

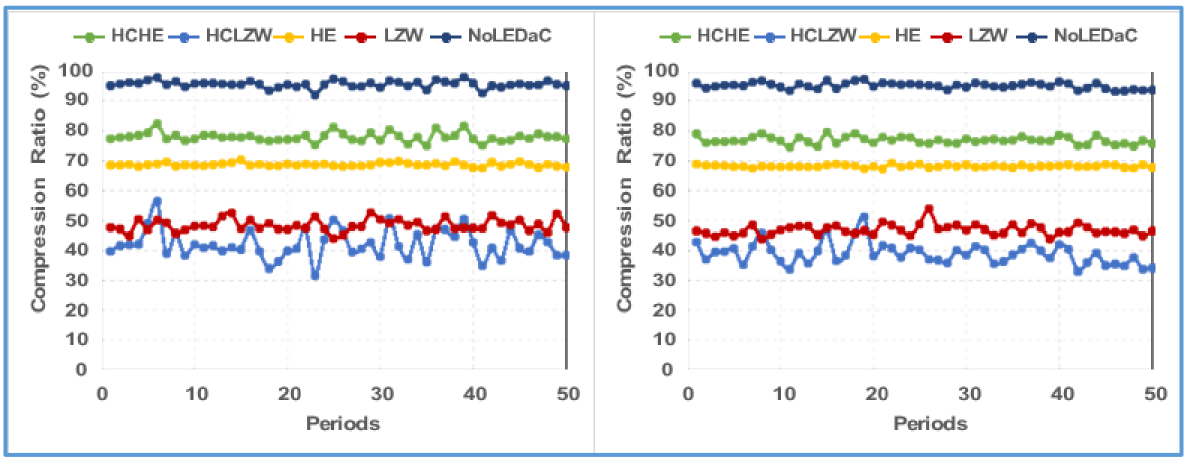
In this section, The original LZW technique, HCLZW, and HCHE are contrasted with our suggested NoLEDaC strategy. The eps parameter is set to 0.20, where eps is the maximum distance between two samples before they are considered to be in close proximity. In addition, the min samples parameter is set to 5, where min samples represent the number of samples (or total weight) in a core point's neighbourhood.

4.5.1 The Compression Ratio

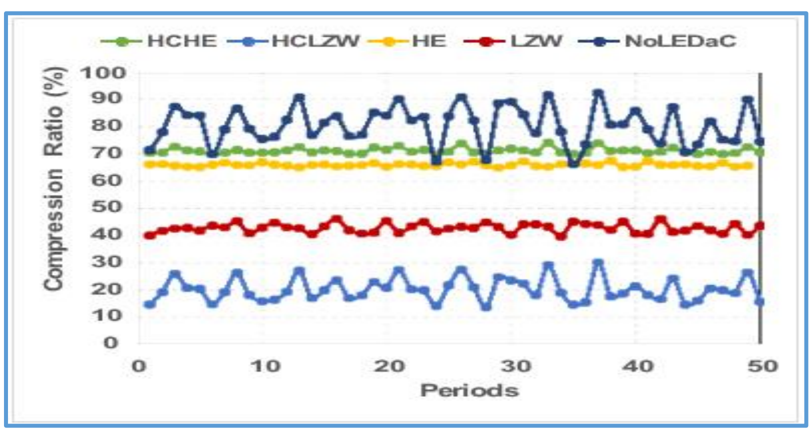
It is a necessary component that affects network performance. This measure is analysed utilising a range of data sources and compares the results derived from various approaches. Figure (4.11) depicts the compression ratio for multiple ways utilising the dataset (N, O, Z, F, S). The data imply that the suggested NoLEDaC methodology can introduce a larger compression ratio than previous techniques. The suggested NoLEDaC method compresses EEG data from 66 percent to 98.6 percent across all records. Whereas the other procedures compress the EEG data from 40% to 54%, from 65% to 70%, from 14% to 56%, and from 70% to 83%, respectively, for all records using LZW, HE, HCLZW, and HCHE.



(a) **(b)**



(c) **(d)**



(e)

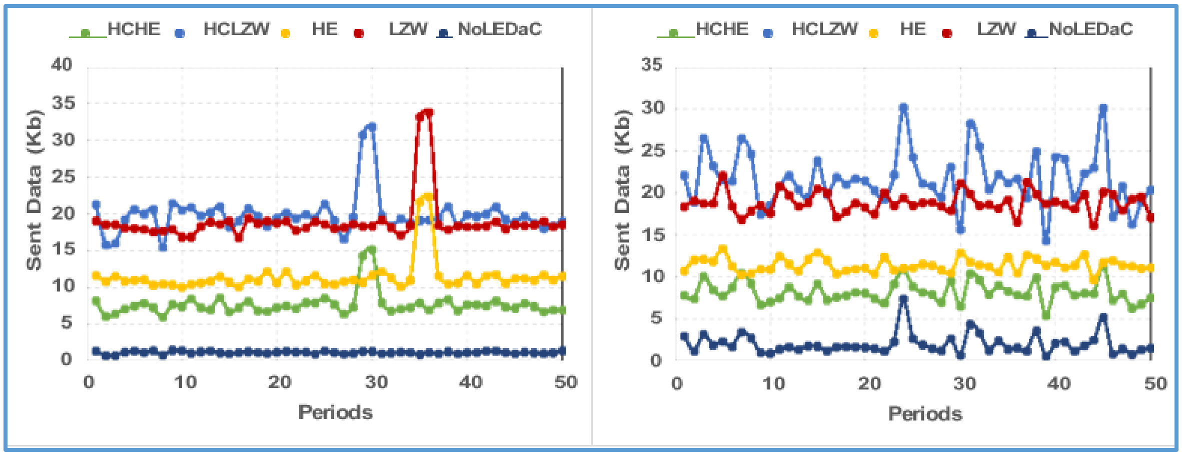
Figure 4.11: The Compression Ratio(NoLEDaC). (a) Z, (b) F, (c) N, (d) O, (e) S.

4.5.2 The Sent EEG Data

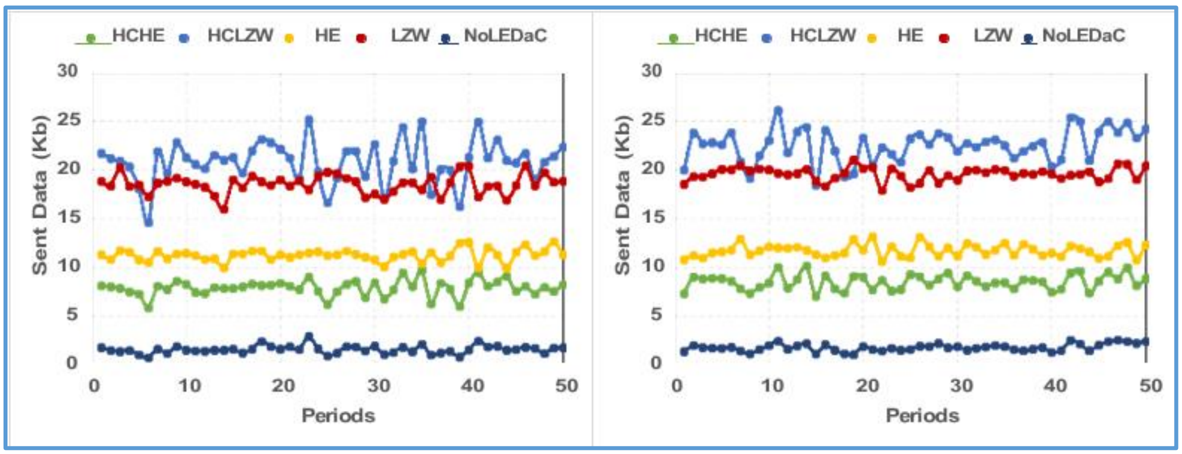
In this experiment, the effect of the proposed NoLEDaC approach on the transfer of EEG data to the cloud server was explored. Figure 4.12 depicts the transmitted data for various ways utilising the dataset (Z, F, N, O, S). The findings indicate that the proposed NoLEDaC approach can provide EEG data in a more compact format than current methods. The results demonstrate that the suggested NoLEDaC method reduces the size of sent EEG data (in KB) by 0.6% to 14% across distinct records. Whereas the other techniques reduced the volume of transmitted EEG data (in KB) by 16 to 34 percent, 10 to 22 percent, 14 to 37 percent, and 5 to 15 percent, respectively, for LZW, HE, HCLZW, and HCHE, for different records.

4.5.3 The Compression Time

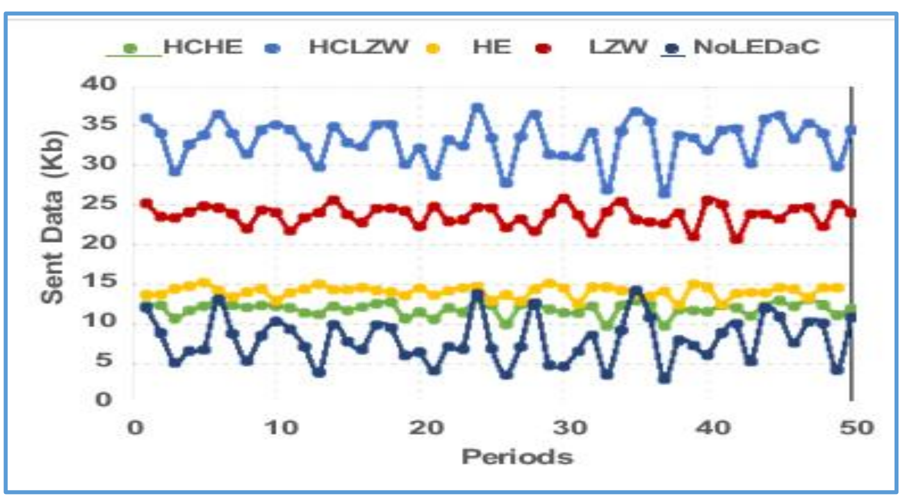
In this experiment, the effect of the proposed NoLEDaC on the compression time of EEG data was explored. Figure 4.13 depicts the compression time for a variety of techniques employing different datasets (Z, F, N, O, S). Compared to HCHE, HCLZW, and LZW, the NoLEDaC method can provide less compression time per period while requiring somewhat more compression time than HE. The proposed NoLEDaC method enhances network performance by decreasing the quantity of transmitted EEG data and boosting the compression ratio, both of which have a substantial impact on network performance.



(a) (b)

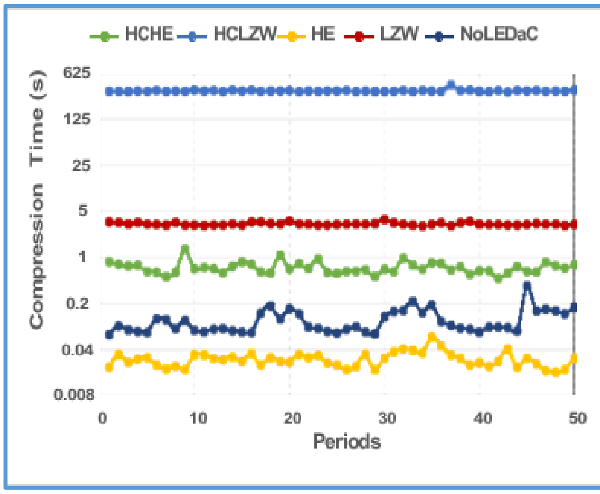


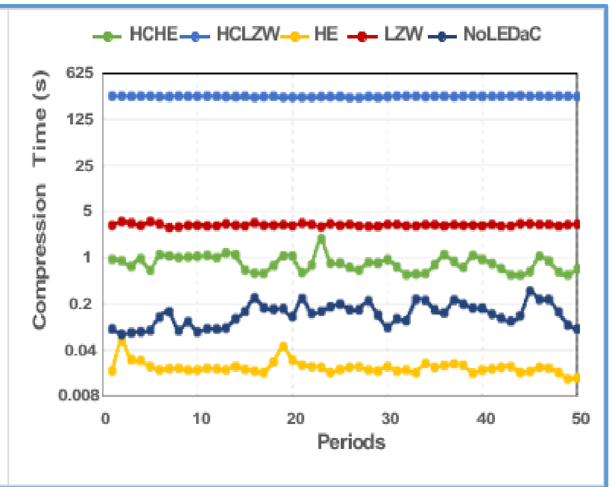
(c) (d)



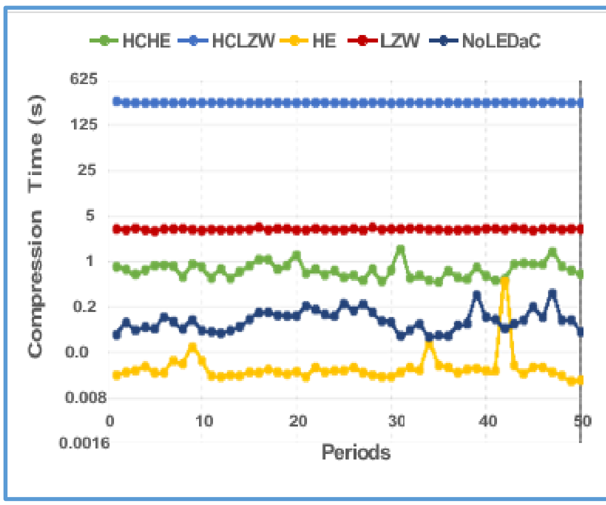
(e)

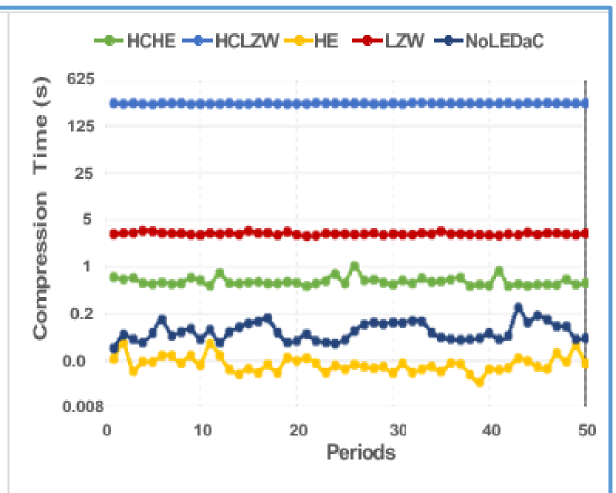
Figure 4.12: The Sent EEG Data(NoLEDaC). (a) Z, (b) F, (c) N, (d) O, (e) S.



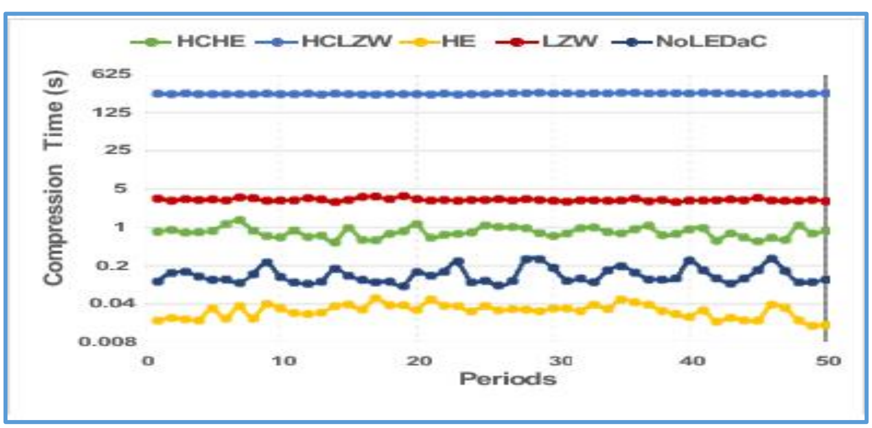


(a) **(b)**





(c) **(d)**

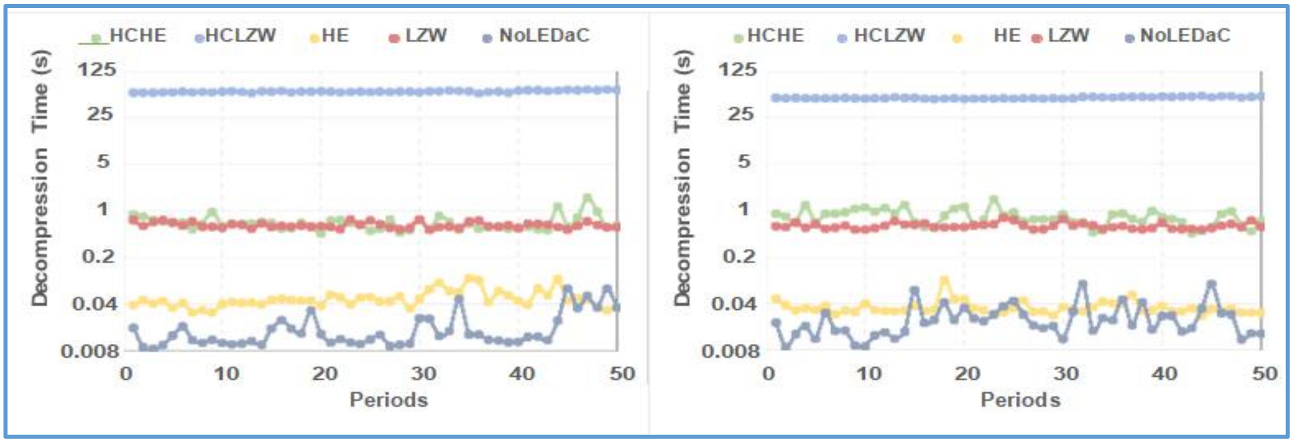


(e)

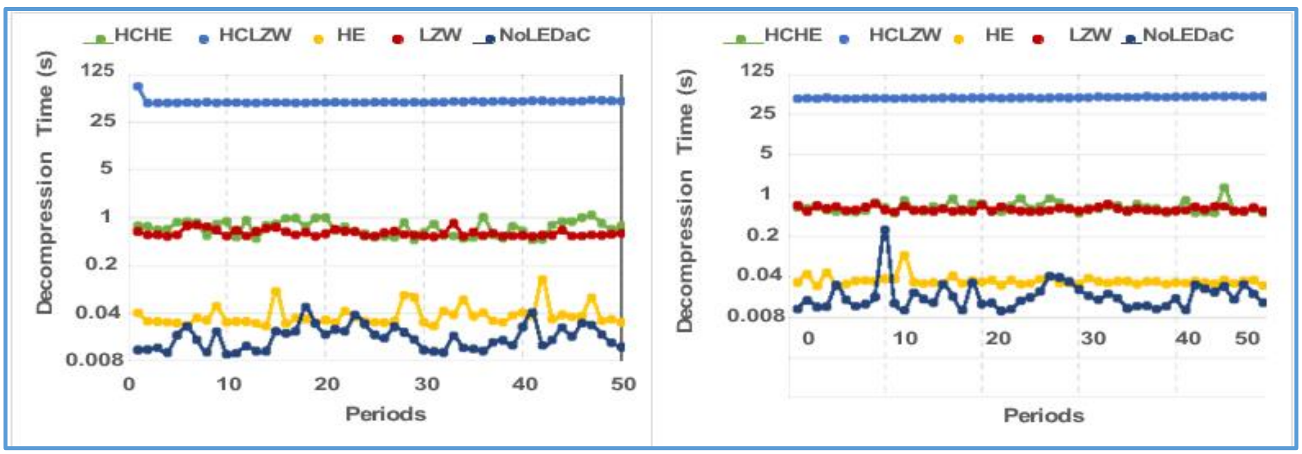
Figure 4.13:The Compression Time(NoLEDaC). (a) Z, (b) F, (c) N, (d) O, (e) S.

4.5.4 The Decompression Time

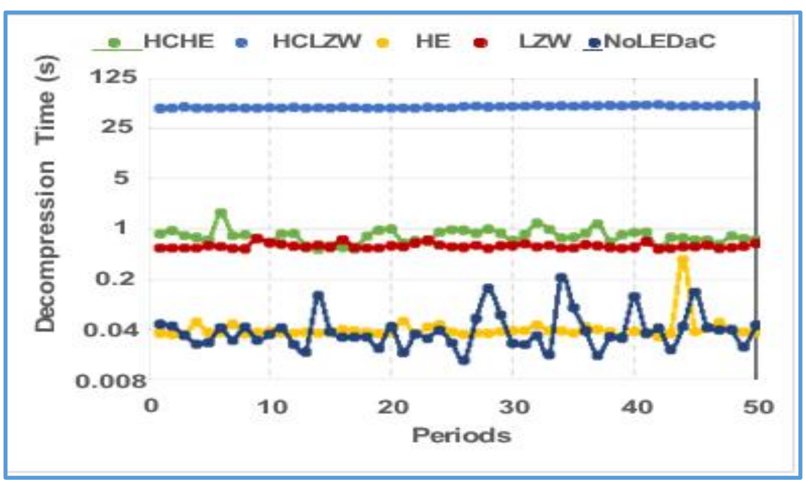
In this section, the impact of the proposed NoLEDaC on the decompression time of "EEG data" was explored. Figure 4.14 depicts the decompression time for many techniques utilizing numerous datasets (N, O, Z, F, S). In comparison to HCHE, HCLZW, HE, and LZW, the results indicate that the NoLEDaC technique can reduce decompression time in the majority of cases while taking slightly more time than HE in some cases.



(a) (b)



(c) (d)



(e)

Figure 4.14:The Decompression Time(NoLEDaC). (a) Z, (b) F, (c) N, (d) O, (e) S.

4.5.6 Compression Power and Average Compression Power

The proposed NoLEDaC is evaluated using the compression power performance metric (see the equations 2.2 and 2.3 in Chapter 2) and compared to several existing compression algorithms in related studies such as 1-D SPIHT [21], 2-D SPIHT + AC [22], 1-D SHORTEN [22], JPEG2000 [22], 2-D SPIHT [21], and AC [22]. Figure 4.15 depicts the compression power of many strategies.

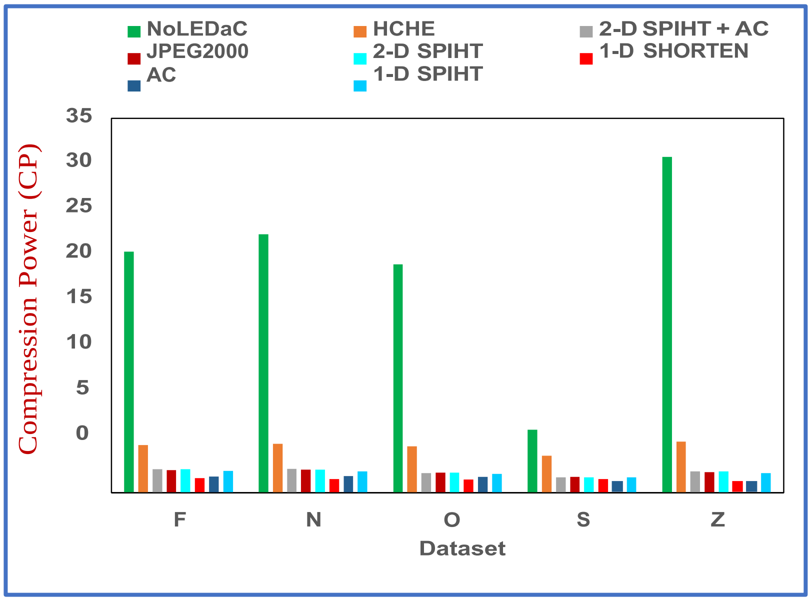


Figure 4.15: The Compression Power(NoLEDaC)

As illustrated in Figure (4.15), the proposed NoLEDaC strategy gives greater performance compared to previous approaches. For various dataset recodings, the compression capability of the proposed NoLEDaC method is between one and six times greater than the compression capability of the existing methods (Z, F, N, O, S).

Additionally, compared to existing systems,The suggested NoLEDaC approach has an average compression power that is more than five times greater (see Figure 4.16). By using the high data similarity among the obtained data at the gateway of the fog layer, the suggested NoLEDaC technique provides a better EEG data size reduction.

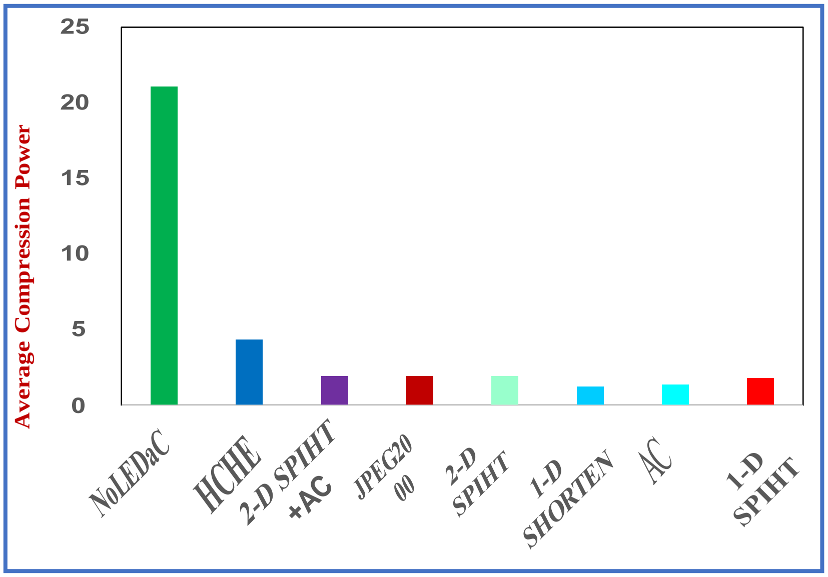


Figure 4.16: Average Compression Power(NoLEDaC)

Chapter Five

Conclusion and future Works

5.1 Conclusions

- In this thesis, three lossless EEG data compression methods are used to reduce data traffic and increase network performance.

- First, a Lossless EEG Data Compression (LEDaC) is proposed using clustering and encoding for (Internet of Medical Things (IoMT) based Fog computing networks.The proposed LEDaC technique introduce a good compression ratio. where it compresses the EEG data per period from 2.1 up to 4.39 for different dataset records (Z, N, S, F, O) and reduce the transmitted data efficiently without losing the data at the receiving side and it transmits from 2.1 KB up to 4.4 KB for different data records also LEDaC technique compress the EEG data of each period with a suitable time. where it requires from 0.95 up to 8.67 seconds, It can seen that the proposed LEDaC technique decompresses the received compressed file at the fog gateway with a suitable time. where it needs from 0.95 up to 7.9 seconds to decompress the received compressed file of one period for different dataset records (Z, N, S, F, O).

- Then, an Efficient Compression Technique (ECoT) is proposed for Reducing Transmitted EEG Data Without Loss in IoMT Networks Based on Fog Computing. The ECoT technique compresses the EEG data from 65% up to 85% for different records also It can be seen from the results that the proposed ECoT reduces the size of sent EEG data (in KB) from 9.4 up to 12.2 for different records. The results indicate that when compared to other approuches the ECoT can give less compression and decompression time per period for different types of records

- Finally, a novel Lossless EEG for Data Compression (NoLEDaC) Approach is proposed at the fog gateway for IoMT.The EEG data is compressed by the proposed NoLEDaC approach from 66% up to 98.6% in all records.also it can be seen from the results that the proposed NoLEDaC approach reduces the size of sent EEG data (in KB) from 0.6% up to 14% in different records. The NoLEDaC approach can give less compression time and decompression time per period, The compression power of the suggested NoLEDaC approach is more than one to six times and The average compression power is more than five times greater than the existing approaches.

- The results demonstrate that in terms of compression ratio, sent data, compression and decompression time, and compression power, the suggested lossless compression approaches perform better than certain relevant studies in the literature.

- The NoLEDaC technique introduce better performance than the LEDaC and ECoT while ECoT outperforms LEDaC technique.

5.2 Future Works

The following future works can be performed:

1. Planning to apply the data compression approach in the smart sensors themselves instead of the fog layer to obtain faster network performance.
2. Planning to use lossy compression approaches to further compress the EEG data while keeping a good data quality.
3. Planning to propose a hybrid compression approach that can use the lossy compression method in the case of normal case of the patient while it must use the lossless compression when the patient in risk.

References

[1] Harb, H., Idrees, A. K., Jaber, A., Makhoul, A., Zahwe, O., & Taam, M. A. (2017). Wireless sensor networks: a big data source in internet of things. *International Journal of Sensors Wireless Communications and Control*, *7*(2), 93-109.

[2] Pattar, S., Buyya, R., Venugopal, K. R., Iyengar, S. S., & Patnaik, L. M. (2018). Searching for the IoT resources: Fundamentals, requirements, comprehensive review, and future directions. *IEEE Communications Surveys & Tutorials*, *20*(3), 2101-2132.

[3] Idrees, A. K., Al-Qurabat, A. K. M., Abou Jaoude, C., & Al-Yaseen, W. L. (2019, June). Integrated divide and conquer with enhanced k-means technique for energy-saving data aggregation in wireless sensor networks. In *2019 15th International wireless communications & mobile computing conference (IWCMC)* (pp. 973-978). IEEE.

[4] Jaber, A. S., & Idrees, A. K. (2021). Energy-saving multisensor data sampling and fusion with decision-making for monitoring health risk using WBSNs. *Software: Practice and Experience*, *51*(2), 271-293.

[5] Papageorgiou, A., Cheng, B., & Kovacs, E. (2015, November). Real-time data reduction at the network edge of Internet-of-Things systems. In *2015 11th international conference on network and service management (CNSM)* (pp. 284-291). IEEE.

[6] Cicioğlu, M., & Çalhan, A. (2021). Energy efficiency solutions for IEEE 802.15. 6 based wireless body sensor networks. *Wireless Personal Communications*, *119*(2), 1499-1513.

[7] Mavinkattimath, S. G., Khanai, R., & Torse, D. A. (2019, April). A survey on secured wireless body sensor networks. In *2019 International Conference on Communication and Signal Processing (ICCSP)* (pp. 0872-0875). IEEE.

[8] Al-Janabi, S., Al-Shourbaji, I., Shojafar, M., & Shamshirband, S. (2017). Survey of main challenges (security and privacy) in wireless body area networks for healthcare applications. *Egyptian Informatics Journal*, *18*(2), 113-122.

[9] Khan,J.Y& Mehmet ,R. Y. (2004). Wireless Body area network for medical applications. Information Tech. In Biomedicine, 8(4), (1089-771) IEEE.

[10] Movassaghi, S., Abolhasan, M., Lipman, J., Smith, D., & Jamalipour, A. (2014). Wireless body area networks: A survey. *IEEE Communications surveys & tutorials*, *16*(3), 1658-1686.

[11] Hayajneh, T., Almashaqbeh, G., Ullah, S., & Vasilakos, A. V. (2014). A survey of wireless technologies coexistence in WBAN: analysis and open research issues. *Wireless Networks*, *20*(8), 2165-2199.

[12] Darwish, A., & Hassanien, A. E. (2011). Wearable and implantable wireless sensor network solutions for healthcare monitoring. *Sensors*, *11*(6), 5561-5595.

[13] Hussain, S., Ullah, S. S., Uddin, M., Iqbal, J., & Chen, C. L. (2022). A comprehensive survey on signcryption security mechanisms in wireless body area networks. *Sensors*, *22*(3), 1072.

[14] Sangwan, A., & Bhattacharya, P. P. (2015). Wireless body sensor networks: a review. *International Journal of Hybrid Information Technology*, *8*(9), 105-120.

[15]Niedermeyer,E.,&daSilva,F.L.(Eds.).(2005). *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins.

[16] Fauvel, S., & Ward, R. K. (2014). An energy efficient compressed sensing framework for the compression of electroencephalogram signals. *Sensors*, *14*(1), 1474-1496.

[17] Seeck, M., Koessler, L., Bast, T., Leijten, F., Michel, C., Baumgartner, C., ... & Beniczky, S. (2017). The standardized EEG electrode array of the IFCN. *Clinical neurophysiology*, *128*(10), 2070-2077.

[18] Troy M, L., Joseph T, G., & Daniel P, F. (2012). How many electrodes are really needed for EEG-based mobile brain imaging?. *Journal of Behavioral and Brain Science*, *2012*.

[19] Abdellatif, A. A., Emam, A., Chiasserini, C. F., Mohamed, A., Jaoua, A., & Ward, R. (2019). Edge-based compression and classification for smart healthcare systems: Concept, implementation and evaluation. *Expert Systems with Applications*, *117*, 1-14.

[20] Abdellatif, A. A., Mohamed, A., & Chiasserini, C. F. (2018). User-centric networks selection with adaptive data compression for smart health. *IEEE Systems Journal*, *12*(4), 3618-3628.

[21] Srinivasan, K., & Reddy, M. R. (2010). Efficient preprocessing technique for real-time lossless EEG compression. *Electronics Letters*, *46*(1), 26-27.

[22] Srinivasan, K., Dauwels, J., & Reddy, M. R. (2011). A two-dimensional approach for lossless EEG compression. *Biomedical signal processing and control*, *6*(4), 387-394.

[23] Maazouz, M., Kebir, S. T., Bengherbia, B., Toubal, A., Batel, N., & Bahri, N. (2015, December). A DCT-based algorithm for multi-channel near-lossless EEG

compression. In *2015 4th International Conference on Electrical Engineering (ICEE)* (pp. 1-5). IEEE.

[24] Birvinskas, D., Jusas, V., Martisius, I., & Damasevicius, R. (2015). Fast DCT algorithms for EEG data compression in embedded systems. *Computer Science and Information Systems*, *12*(1), 49-62.

[25] Karimu, R. Y., & Azadi, S. (2016). Lossless EEG compression using the DCT and the Huffman coding.

[26] Alsenwi, M., Ismail, T., & Mostafa, H. (2016, December). Performance analysis of hybrid lossy/lossless compression techniques for EEG data. In *2016 28th International Conference on Microelectronics (ICM)* (pp. 1-4). IEEE.

[27] Hejrati, B., Fathi, A., & Abdali-Mohammadi, F. (2017). Efficient lossless multi-channel EEG compression based on channel clustering. *Biomedical Signal Processing and Control*, *31*, 295-300.

[28] Alsenwi, M., Saeed, M., Ismail, T., Mostafa, H., & Gabran, S. (2017, December). Hybrid compression technique with data segmentation for electroencephalography data. In *2017 29th International Conference on Microelectronics (ICM)* (pp. 1-4). IEEE.

[29] Azar, J., Darazi, R., Habib, C., Makhoul, A., & Demerjian, J. (2018, June). Using dwt lifting scheme for lossless data compression in wireless body sensor networks. In *2018 14th International wireless communications & mobile computing conference (IWCMC)* (pp. 1465-1470). IEEE.

[30] Al-Nassrawy, K. K., Al-Shammary, D., & Idrees, A. K. (2020). High performance fractal compression for EEG health network traffic. *Procedia Computer Science*, *167*, 1240-1249.

[31] Rajasekar, P., & Pushpalatha, M. (2020). Huffman quantization approach for optimized EEG signal compression with transformation technique. *Soft Computing*, *24*(19), 14545-14559.

[32] Das, S., & Kyal, C. (2021). Efficient multichannel EEG compression by optimal tensor truncation. *Biomedical Signal Processing and Control*, *68*, 102749.

[33] Idrees, S. K., & Idrees, A. K. (2022). New fog computing enabled lossless EEG data compression scheme in IoT networks. *Journal of Ambient Intelligence and Humanized Computing*, *13*(6), 3257-3270.

[34] Campobello, G., Gugliandolo, G., & Donato, N. (2021, August). A Simple and Efficient Near-lossless Compression Algorithm for Multichannel EEG Systems. In *2021 29th European Signal Processing Conference (EUSIPCO)* (pp. 1150-1154). IEEE.

[35] Al-Nassrawy, K. K., Idrees, A. K., & Al-Shammary, D. (2022). A Novel Lossless EEG Compression Model Using Fractal Combined with Fixed-Length Encoding Technique. In *AI and IoT for Sustainable Development in Emerging Countries* (pp. 439-454). Springer, Cham.

[36] Idrees, A. K., Idrees, S. K., Couturier, R., & Ali-Yahiya, T. (2022). An Edge-Fog Computing Enabled Lossless EEG Data Compression with Epileptic Seizure Detection in IoMT Networks. *IEEE Internet of Things Journal*.

[37] Gope, P., & Hwang, T. (2015). Untraceable sensor movement in distributed IoT infrastructure. *IEEE Sensors Journal*, *15*(9), 5340-5348.

[38] Gope, P., & Hwang, T. (2015). A realistic lightweight authentication protocol preserving strong anonymity for securing RFID system. *computers & security*, *55*, 271-280.

[39] Alkhayyat, A., Thabit, A. A., Al-Mayali, F. A., & Abbasi, Q. H. (2019). WBSN in IoT health-based application: toward delay and energy consumption minimization. *Journal of Sensors*, *2019*.

[40] Latré, B., Braem, B., Moerman, I., Blondia, C., & Demeester, P. (2011). A survey on wireless body area networks. *Wireless networks*, *17*(1), 1-18.

[41] Bharadwaj, H. K., Agarwal, A., Chamola, V., Lakkaniga, N. R., Hassija, V., Guizani, M., & Sikdar, B. (2021). A review on the role of machine learning in enabling IoT based healthcare applications. *IEEE Access*, *9*, 38859-38890.

[42] Chaudhary, S., Singh, A., & Chatterjee, K. (2019). Wireless body sensor network (wbsn) security and privacy issues: A survey. *International Journal of Computational Intelligence & IoT*, *2*(2).

[43] Rajput, A., & Brahimi, T. (2019). Characterizing IOMT/personal area networks landscape. *arXiv preprint arXiv:1902.00675*.

[44] Bickford, R.(1987) .Electroencephalography. In: Adelman, G. (Ed.), Encyclopedia of Neuroscience, Cambridge, pp. 371–373 .

[45] Demos, J. N. (2005). *Getting started with neurofeedback*. WW Norton & Company.

[46] Foldvary-Schaefer, N., & Grigg-Damberger, M. M. (2012). Identifying interictal and ictal epileptic activity in polysomnograms. *Sleep Medicine Clinics*, *7*(1), 39-58.

[47] Hussain, S., Ullah, S. S., Uddin, M., Iqbal, J., & Chen, C. L. (2022). A comprehensive survey on signcryption security mechanisms in wireless body area networks. *Sensors*, *22*(3), 1072.

[48] Zimmerman, T. G. (1996). Personal area networks: Near-field intrabody communication. *IBM systems Journal*, *35*(3.4), 609-617.

[49] Barakah, D. M., & Ammad-uddin, M. (2012, February). A survey of challenges and applications of wireless body area network (WBAN) and role of a virtual doctor server in existing architecture. In *2012 Third International Conference on Intelligent Systems Modelling and Simulation* (pp. 214-219). IEEE.

[50] Mavinkattimath, S. G., Khanai, R., & Torse, D. A. (2019, April). A survey on secured wireless body sensor networks. In *2019 International Conference on Communication and Signal Processing (ICCSP)* (pp. 0872-0875). IEEE.

[51] Garg, M., & Gupta, M. (2015). Technical issues and challenges in building human body sensor networks. *International Journal of Advanced Computer Science and Applications*, *6*(6).

[52] Chen, C. A., Chen, S. L., Huang, H. Y., & Luo, C. H. (2011). An asynchronous multi-sensor micro control unit for wireless body sensor networks (WBSNs). *Sensors*, *11*(7), 7022-7036.

[53] Roy, N. B., & Das, D. (2015). Role of body area sensor networks in smart health care. *International Journal of Future Computer and Communication*, *4*(5), 320.

[54] Fu, L., & Pan, M. (2009, October). A simplified FPGA implementation based on an Improved DES algorithm. In *2009 Third International Conference on Genetic and Evolutionary Computing* (pp. 227-230). IEEE.

[55] Chuang, H. C., Shih, C. Y., Chou, C. H., Huang, J. T., & Wu, C. J. (2014). The development of a blood leakage monitoring system for the applications in hemodialysis therapy. *IEEE Sensors Journal*, *15*(3), 1515-1522.

[56] Shirazi, S. N., Gouglidis, A., Farshad, A., & Hutchison, D. (2017). The extended cloud: Review and analysis of mobile edge computing and fog from a security and resilience perspective. *IEEE Journal on Selected Areas in Communications*, *35*(11), 2586-2595.

[57] Muniswamaiah, M., Agerwala, T., & Tappert, C. C. (2021, June). Fog computing and the internet of things (IoT): a review. In *2021 8th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/2021 7th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom)* (pp. 10-12). IEEE.

[58] Sabireen, H., & Neelanarayanan, V. (2021). A review on fog computing: Architecture, fog with IoT, algorithms and research challenges. *Ict Express*, *7*(2), 162-176.

[59] Kumar, K. M., & Reddy, A. R. M. (2016). A fast DBSCAN clustering algorithm by accelerating neighbor searching using Groups method. *Pattern Recognition*, *58*, 39-48.

[60] Ghosal, A., Nandy, A., Das, A. K., Goswami, S., & Panday, M. (2020). A short review on different clustering techniques and their applications. *Emerging technology in modelling and graphics*, 69-83.

[61] Fahad, A., Alshatri, N., Tari, Z., Alamri, A., Khalil, I., Zomaya, A. Y., ... & Bouras, A. (2014). A survey of clustering algorithms for big data: Taxonomy and empirical analysis. *IEEE transactions on emerging topics in computing*, *2*(3), 267-279.

[62] Khan, K., Rehman, S. U., Aziz, K., Fong, S., & Sarasvady, S. (2014, February). DBSCAN: Past, present and future. In *The fifth international*

conference on the applications of digital information and web technologies (ICADIWT 2014) (pp. 232-238). IEEE.

[63] Agrawal, R., Gehrke, J., Gunopulos, D., & Raghavan, P. (1998, June). Automatic subspace clustering of high dimensional data for data mining applications. In *Proceedings of the 1998 ACM SIGMOD international conference on Management of data* (pp. 94-105).

[64] Bhatia, P. (2019). Data Mining and Data Warehousing: Principles and Practical Techniques. Cambridge: Cambridge University Press. doi:10.1017/9781108635592

[65] Pandey, M., Shrivastava, S., Pandey, S., & Shridevi, S. (2020, February). An enhanced data compression algorithm. In *2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)* (pp. 1-4). IEEE.

[66] Jayasankar, U., Thirumal, V., & Ponnurangam, D. (2021). A survey on data compression techniques: From the perspective of data quality, coding schemes, data type and applications. *Journal of King Saud University-Computer and Information Sciences*, *33*(2), 119-140.

[67] Akhter, S., & Haque, M. A. (2010, August). ECG comptression using run length encoding. In *2010 18th European Signal Processing Conference* (pp. 1645-1649). IEEE.

[68] Drweesh, Z. T., & George, L. E. (2014). Audio compression based on discrete cosine transform, run length and high order shift encoding. *International Journal of Engineering and Innovative Technology (IJEIT)*, *4*(1), 45-51.

[69] Saidani, A., Jianwen, X., & Mansouri, D. (2020). A lossless compression approach based on delta encoding and T-RLE in WSNs. *Wireless Communications and Mobile Computing*, *2020*.

[70] Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N., & Wolpaw, J. R. (2004). BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Transactions on biomedical engineering*, *51*(6), 1034-1043.

[71] DelPozo-Banos, M., Travieso, C. M., Weidemann, C. T., & Alonso, J. B. (2015). EEG biometric identification: a thorough exploration of the time-frequency domain. *Journal of neural engineering*, *12*(5), 056019.

[72] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.

[73] Bao, X., Wang, J., & Hu, J. (2009, June). Method of individual identification based on electroencephalogram analysis. In *2009 international conference on new trends in information and service science* (pp. 390-393). IEEE.

[74] Brunner, C., Leeb, R., Müller-Putz, G., Schlögl, A., & Pfurtscheller, G. (2008). BCI Competition 2008–Graz data set A. *Institute for Knowledge Discovery (Laboratory of Brain-Computer Interfaces), Graz University of Technology*, *16*, 1-6.

[75] Lawhern, V. J., Solon, A. J., Waytowich, N. R., Gordon, S. M., Hung, C. P., & Lance, B. J. (2018). EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces. *Journal of neural engineering*, *15*(5), 056013.

[76] Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual

complexity. *Journal of experimental psychology: Human learning and memory*, *6*(2), 174.

[77] Liew, S. H., Choo, Y. H., Low, Y. F., & Yusoh, Z. I. M. (2015). Identifying visual evoked potential (VEP) electrodes setting for person authentication. *Int. J. Adv. Soft Comput. Appl*, *7*(3), 85-99.

[78] Andrzejak, R. G., Lehnertz, K., Mormann, F., Rieke, C., David, P., & Elger, C. E. (2001). Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. *Physical Review E*, *64*(6), 061907.

[79] Dhull, S. K., & Singh, K. K. (2021). A review on automatic epilepsy detection from EEG signals. *Advances in Communication and Computational Technology*, 1441-1454.

Publications

1. Marwa saieed khlief, Ali Kadhum Idrees. "Efficient EEG Data Compression Technique for Internet of Health Things Networks“, Accepted in 2022 IEEE World Conference on Applied Intelligence and Computing (AIC 2022) to be held on June 17-19, 2022.

2. Marwa saieed khlief, Ali Kadhum Idrees. A Comprehensive Review of the EEG Data Analytics, International Journal of Computer Applications in Technology, Clarivate, Accepted.

3. Marwa saieed khlief, Ali Kadhum Idrees. Lossless EEG Data Compression using Clustering and Encoding for Fog Computing based IoMT Networks, International Journal of Computer Applications in Technology, Clarivate, Accepted.

4. Marwa saieed khlief, Ali Kadhum Idrees. A New Lossless EEG Data Compression Approach Enabled by Fog Computing in Internet of Health Things Networks, Submitted to Computational intelligence, Under review.

5. Marwa saieed khlief, Ali Kadhum Idrees. Efficient Compression Technique for Reducing Transmitted EEG Data Without Loss in IoMT Networks Based on Fog Computing, submitted Int. J. of Ad Hoc and Ubiquitous Computing, Under review.

الملخص

ظهرت إنترنت الأشياء الطبية (IoMT) ، والتي تضم العديد من التطبيقات المبتكرة التي ترتبط بأنظمة الرعاية الصحية من خلال شبكات الكمبيوتر عبر الإنترنت بسبب النمو السريع للمعدات الطبية وتكنولوجيا الاتصالات. حيث ستنشئ هذه التطبيقات بانتظام كميات هائلة من البيانات ، والتي ستنقلها بوابة الحافة إلى السحابة بصورة مستمرة ومنتظمه لمزيد من المعالجة.حيث سيتم فرض ضرائب كبيرة على شبكة إنترنت الأشياء إذا تم إرسال هذا الحجم الهائل من البيانات عبرها إلى السحابة. تؤثر كمية البيانات المرسلة والتأخيرات المطولة في المعالجة بشكل كبير على سرعة تفاعل تطبيقات إنترنت الأشياء. وبذلك سوف يتم تقليل أوقات استجابة هذه التطبيقات. ولتقليل كمية البيانات المنقولة وتحسين أوقات رد الفعل سوف تستفيد تطبيقات إنترنت الأشياء من مزايا الحوسبة الضبابية ، والتي تعمل كمستوى وسيط بين الاجهزة الذكية والإنترنت.

تقترح الأطروحة الحالية ضغط بيانات EEG بدون فقدان لتطبيقات الحوسبة الضبابية لانترنت الاشياء حيث تم توفير ثلاث طرق للتشفير يمكن إجراؤها عند بوابة الضباب لتقليل نقل بيانات EEG الخاصة بالمريض إلى مركز البيانات السحابية وهذة الطرق هي : أولاً , تقنية ضغط بيانات EEG بدون فقدان (LEDaC) من أجل تقليل كمية بيانات IoMT ، تجمع (LEDaC) بين تقنيتين فعالتين لتجميع DBSCAN وتشفير Huffmanبدلاً من تحميل البيانات للمعالجة والتحليل الإضافي الى مركز البيانات السحابي. ثانيا , تم اقتراح تقنية ضغط فعالة (ECoT) لضغط (EEG) دون خسارة. من أجل توصيل بيانات EEG الضرورية فقط إلى السحابة ،حيث تدمج تقنية(EcoT) ثلاث طرق فعالة لتقليل البيانات في بوابة الضباب وهي: DBSCAN Clustering، وتشفير Delta وتشفير Huffman. ثالثا ، تم اقتراح تقنية (NoLEDaC) لضغط بيانات EEG في الحوسبة الضبابية لشبكات إنترنت الأشياء الطبية. حيث تقلل هذة الطريقة من حجم بيانات EEG الخاصة بالمريض عن طريق ضغطها في عقدة الضباب قبل تحميلها إلى مركز البيانات السحابي. تجمع طريقة (NoLEDaC) المقترحة بين: DBSCAN Clustering متبوعة بطريقة الضغط RLE وضغط Huffman عند بوابة الضباب.

تم إكمال العديد من التجارب بنجاح باستخدام مجموعة بيانات (Bonn) أظهرت النتائج أن تقنيات الضغط غير المنقوص المقترحة تتفوق على التقنيات الأخرى من حيث نسبة الضغط ووقت الضغط ووقت فك الضغط وحجم البيانات المرسلة وقوة الضغط ومتوسط قوة الضغط. تقدم تقنية LEDaC المقترحة نسبة ضغط جيدة ، حيث تقوم بضغط بيانات EEG لكل فترة من 2.1 حتى 4.39 لسجلات مجموعة بيانات مختلفة ، بينما تقوم تقنية ECoT بضغط بيانات EEG من 65٪ إلى 85٪ لسجلات مختلفة. تم ضغط بيانات EEG بواسطة تقنية NoLEDaC والحصول على نسبة ضغط أعلى من ECoT و LEDaC بمعدل من 66٪ إلى 98.6٪ في جميع السجلات.

وزارة التعليم العالي و البحث العلمي

جامعة بابل كلية العلوم للبنات

قسم علوم الحاسوب

اسلوب ضغط بيانات مخطط كهربائية الدماغ الفعال لتطبيقات انترنت الاشياء القائمة على الحوسبة الضبابية

رسالة مقدمة الى مجلس كلية العلوم للبنات في جامعة بابل وهي جزء من متطلبات الحصول على درجة الماجستير في علوم الحاسبات

مقدمة من قبل

مروة سعيد خليف

باشراف

الاستاذ الدكتور

علي كاظم ادريس السعدي

1444 هـ **2022 م**