International Journal of Communication Networks and Information Security

ISSN: 2073-607X, 2076-0930

*Volume **14** Issue **03** Year **2022** Page **313:329***

5G with Fog Computing based Privacy System in Data Analytics for Healthcare System by AI Techniques

Encarnación Maria Torres Solís

Department of Post Grade, Universidad Cèsar Vallejo, Perú, South America

Juan Carlos Cotrina-Aliaga

Faculty of Human Medicine, Universidad Privada San Juan Bautista, Chincha, Perú, South America

juan.cotrina@upsjb.edu.pe

Fidel Castro-Cayllahua

Associate Professor, Universidad Peruana los Andes, Perú, South America

Dr. Severo Simeón Calderón Samaniego

Principal Professor, Universidad Peruana los Andes, Perú, South America

Betsy Nordie Pardo Alarcon

Department of Post Grade, Universidad Cèsar Vallejo, Perù, South America

Yoni Magali Maita Cruz

Professional School of Human Medicine, Universidad Privada San Juan Bautista, Perú, South America

Article History

Received: 13 July 2022
Revised: 20 September 2022
Accepted: 26 October 2022

Abstract

Fog computing architecture is an extended version of the cloud computing architecture to reduce the load of the data transmission and storage in the cloud platform. The architecture of the fog increases the performance with improved efficiency compared with the cloud environment. The fog computing architecture uses the 5G based Artificial Intelligence (AI) technology for performance enhancement. However, due to vast range of data availability privacy is challenging in the fog environment. This paper proposed a Medical Fog Computing Load Scheduling (MFCLS) model for data privacy enhancement. The developed architecture model of optimization-based delay scheduling for task assignment in the fog architecture. The healthcare data were collected and processed with the 5G technology. The developed MFCLS model uses the entropy-based feature selection for the healthcare data. The proposed MFCLS considers the total attributes of 13 for the evaluation of features. With the provision of service level violation, the fog computing network architecture will be provided with reduced energy consumption. The developed load balancing reduced the service violation count with the provision of desired data privacy in the fog model. The estimation of the time frame is minimal for the proposed MFCLS model compared with the existing DAG model. The performance analysis expressed that SLRVM and ECRVM achieved by the proposed MFCLS are 28 and 43 respectively. The comparative examination of the proposed MFCLS model with the existing DAG model expressed that the proposed model exhibits ~6% performance enhancement in the data privacy for the healthcare data.

Keywords: Fog computing, Load balancing, Optimization, data privacy, energy consumption

1. Introduction

The term healthcare refers to a system that entails the enhancement of health-related services in order to fulfill the clinical requests of the individuals. In healthcare services, patients, doctors, clinicians, researchers and medical industries are all making an effort to maintain and restore health records [1]. In recent years, with the remarkable development of technologies, data is continuously increasing day by day in every sector including healthcare which in turn demands more and more data mining applications. However, due to digitization of healthcare system, the medical organizations are generating large amount of healthcare data [2]. In general, healthcare data consists of all records related to health stored in a digital form. It may contain detailed information about patients' medical history, doctors prescribed notes, clinical reports etc. All these data are voluminous, high dimensional and diversified in nature. Healthy decision making is a challenge in the modern era due to this rising complexity of healthcare data [3]. Machine learning, data mining, and statistical techniques are the key fields of study that enhance the ability of individuals to make the right decisions in order to maximize the outcome of any working domain [3]. Human data analytical capability rate is much smaller when compared to the amount of data that is stored [4]. This becomes even more critical when it comes to healthcare domain as the number of available experts for healthcare data analysis is comparatively less.
With the extension of the Cloud computing (CC) Cisco presented a Fog computing technology to meet the requirement of the present era requirement such as 5G, Internet of Things, Artificial Intelligence (AI) and so on [5]. The provision of services by the Fog computing (FC) comprises of IoT users to perform data processing and storage. The technology of Fog computing is extended towards the cloud computing technology for the stored data in the local node which increases the burden towards the data transmission and reception in the cloud [6]. In this manner, Fog computing increases the efficiency of the cloud network performance. Additionally, with the fog computing the data processing rate in the cloud is reduced with the transfer of information in the cloud storage and analysis of the data. The provision of services to the fog computing architecture decreases the latency and network traffic in the environment. The position of the fog computing relies on the fog node over the complete network [7]. The fog computing architecture comprises of the controller devices which act as a switch to the deployed fog nodes for the targeted area in the environment. The data is generated with the use of IoT devices involved in the analysis of information in the nodes. The fog node analyses the data those are transmitted in the cloud with the minimized pressure in the cloud [8]. The reduction of burden in the cloud minimizes the latency and overcrowding in the cloud jobs. The architecture of the cloud computing comprises of the fog computing environment to provide information in the centralized resource access. In other hand, with the decentralized manner fog computing is estimated for the provision of access.
Fog computing architecture exhibits the privileged cloud computing environment to increases the effectiveness and efficiency of the network with the increased number of devices in the network [9]. The fog computing technology comprises of the innovation concept to evaluate the IoT devices storage space and computing ability. The conventional fog computing technology for the medical application comprises of the segments in three tier manner such as IoT Devices (Lower Tier) - responsible for patient information collection, Fog- nodes (Middle Tier), and Cloud (Top Tier). The IoT layer of the fog comprises of the objects those are wearable devices for the data collection such as cameras, vehicles, sensors, home appliances those leads to massive collection of data [10]. The cloud environment provides the assistance in the foundation level for the information upholds to evaluate the framework. However, the vast range of applications demands for the reduced latency to improve the performance of the fog [11]. In reduction of the latency in the fog computing environment the services are reduced with the overcrowding in the network. To evaluate the massive volume of the data enormous data need to be processed and handled with the effective storage information handling in the dynamic environment for the broadband information frames. In IoT application the distributed cloud computing environment demands for the data privacy for the cloud storage those are capable to achieve the effective Quality of Service (QoS) based on the demand [12]. The provision data privacy leads to higher computational capability with the lot of service operation to evaluate large data

volumes. The conventional IoT application subjected to different challenges due to lack of network capacity, latency, location awareness and support. To withstand the issues related to fog computing cloud solution need to be leverages to provide large scale connectivity in the IoT devices to ensure significant solution constraints [13]. With the virtualized platform in fog computing comprises of the resource pool for the business model storage, end-user networking and cloud data centres.
This paper developed a medical data fog architecture model for the task scheduling in the network with the task assignment. The developed model involved in the estimation of the entropy features with the job scheduling and assignment of tasks in the network. This paper is organized as: Section 2 provides the existing literature and proposed methodology presented in section 3. The results obtained for the proposed MFCLS model is presented in section 4 and overall conclusion is presented in section 5.

2. Related Works

Data mining can be defined as a process of collecting, analysing and storing data in order to extract useful and interesting information from it. A few of the key tasks in data mining involve classification, prediction, association analysis, clustering, and hybrid approaches [14]. Notably, data mining is now considered as a multidisciplinary area. In recent years, since the amount of digitally recorded healthcare data has grown, so more and more data mining-based applications are developed by researchers and scientists in order to improve healthcare services. This huge amount of healthcare data can be mined in order to discover new patterns and value-added insights to enhance the quality of healthcare. In the medical field, clinical diagnostic systems are developed to facilitate medical practitioners in making healthcare decisions based on health records. According to [15], electronic health records (EHRs) empower the clinicians by integrating patient health history for planning safe and proper treatment. With a comprehensive evaluation of healthcare medicine based on consideration of six tasks such as prognosis, management, monitoring, screening, and treatment.
In [16] designed a preterm birth model based on the pregnant women dataset those are considered as diabetes mellitus or gestational diabetes mellitus. The logistic regression and Support Vector Machine are used for the prediction purpose. With the developed a multi-parametric magnetic resonance imaging (MRI) machine learning model to predict the risk of breast cancer. In this regard, eight classifiers were used on MRI data of the breast in order to rank the features prediction with the pathologies responsible for treatment of breast cancer. In [17] have introduced a multilevel perceptron model for survival prediction. The introduced model predicts the survival of non-small cell lung cancer patients with multi-layer neural network. Relief and Recursive Feature Elimination algorithms were implemented to extract the features from clinical data in order to increase the accuracy of prediction model. The suggested an intelligent-based expert system for identifying Type2 Diabetes Mellitus (T2DM) patients with the help of electronic health record (EHR) data. For this purpose, different machine learning algorithms such as Random Forest (RF), logistic regression (LR), Decision Tree (DT), naïve bayes, K-Nearest Neighbour (KNN) and Support Vector Machine (SVM) were applied.
In [18] presented a Computer aided diagnosis (CAD) system integrated with the fuzzy k-nearest neighbour (FKNN) classifier to detect and diagnosis thyroid diseases. A thyroid prediction system by employing various machine learning techniques including Decision Trees, Random Forest, Support Vector Machine, ANN and logistic regression. In [19] suggested a heart disease prediction system to detect the presence of cardiovascular disease based on data mining approaches. The three learning classifiers namely, Naïve Bayes, J48 Decision Tree and Bagging algorithms are used in this diagnostic model. According to big data analytics are applied to healthcare data acquired from several sources to gain valuable insight from it. The authors evaluate historical health data and assess medical quality of service from multiple US states by applying new healthcare-specific analytical software to predict future healthcare activities. Despite this, no solutions for storing big data have been developed.
In [20] described the application of data mining techniques in healthcare and biomedicine. The study first examines three different learning methods namely, classification, clustering and association, and then turns its attention to their application in healthcare industry. This is accomplished by a brief discussion given on each task with their advantages and disadvantages. Some of the most important medical aspects that discussed in this paper are: prediction of healthcare costs, disease diagnosis and

prognosis, and discover hidden patterns from healthcare data. However, in most cases, an accurate data mining model is not suitable for clinical environment.

In [21] developed a energy-efficient resource scheduling scheme for the reduction of energy consumption in the IoT based heterogeneous environment. The developed model uses the optimized approach for the data transmission scheduling to reduce latency with the increased network traffic in the network. The simulation analysis expressed that proposed model increases the energy utilization rate by 18% and reduces the overall execution cost of network by 15%. Also, the developed model increases the overall lifetime of the network by 1.17% and 5% increase in the sensor lifetime through the scheduling scheme. In [22] proposed a heuristics method for the scheduling of tasks in the cloud platform. Consequently, the developed model comprises of the effective task scheduling for the complicated procedure based on the consideration of different assumptions. The evaluation is based on the effective task scheduling with the FCFS (First Come First Serve), Minimum Completion Time (MCT), Minimum Execution Time (MET). In [23] developed a task scheduling scheme for the resource allocation in the fog network for the developed application to perform task scheduling and resource redeployment. The analysis expressed that developed model decreases the delay 39% 11% 11% 21% 18% with the increased tasks scheduling in fog nodes.

3. Medical Fog Optimization for Load Balancing

Let $P_c$be the energy consumed for processing instruction in a task and P_t be the energy used for transmission of data. Then choose to perform the task on fog, then the energy consumption is given by the following formula.

$$E_{fog} = P_t + \frac{S}{M} \tag{1}$$

Where S denoted as the byte of data those are processed with the network bandwidth M. Assume the generates output is denoted as S' those are processed effectivey where S' is less than S. Now the smart devices are performed based on the estimation of energy consumption task defined as in equation (2)

$$E_{mobile} = P_c * I + P_t + \frac{S}{M} \tag{2}$$

Where, the task execution in the numerals are denoted as I, the saved energy is represented as (E_s) presented in equation (3) and (4)

$$E_s = P_t * \frac{S}{M} - P_c * I - P_t * \frac{S'}{M} \tag{3}$$

$$E_s = \left[P_t * \frac{S}{M}\right] - P_c * I - [P_t * (S * M)] \tag{4}$$

Where $M = \frac{S'}{s}$and s are called the compression ratio as in equation (5)

$$E_s M = P_t S(1 - M) - P_c * I \tag{5}$$

The saved energy in the Medical Fog Computing Load Scheduling (MFCLS) comprises of the positive evaluation in the task offloading. At first, based on the requrest the CPU is bounded based on the output memory. The Proposed MFCLS algorithm compriseso f the different instances for the classification those are logrithimc algorithm, polynomial algorithm and time. With the fog computing environment the fog computation is based on the reduction of power consumption with the reduced frequency as presented in equation (6)

$$P = cfV^2 + P_{static} \tag{6}$$

In the above equation (6), the c represented as the gate capacitance transistor, frequency of the medica data is denoted as f and the operation voltage and static performance is represented as V. The required voltage for the normal and static operation is based on the frequency dependent with the clocked circuit. With the privacy of the data the energy consumption of the data is denoted as E with the computation of the Power*Time in which E is proportional to V^2 as given in equation (7)

$$E \propto V^2 (V \propto f) \tag{7}$$

The MFCLS algorithm comprises of the scheduling process for job processing based on the consideration of the VM position. The VM is evaluated based on the consideration of the arbitrary process those are applied with the different VM1, VM2 and VM3. The estimation of the arbitrary value comprises of the cost 1 to 3 it reaches till 15. In figure 1 the architecture of the fog computing model is presented.

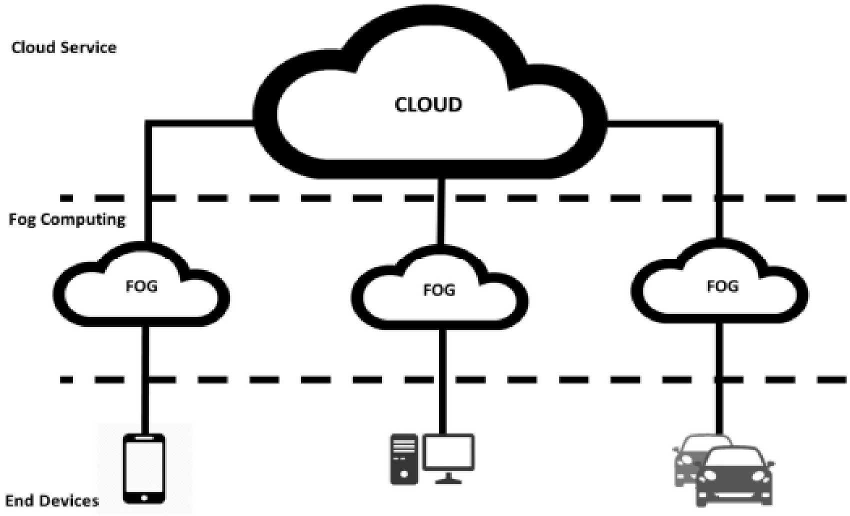

Figure 1. Architecture of Fog Network

If a particular arbitrary cost is taken say from 1 to 3, which are 15.
The structure of job 2 is presented below. Firstly, contemplate the second job:

1. To execute the Job 2 there are 3 options available in the VM.
2. The allocation of Job 2 is based on the obliged VM in the network environment
   a. The slot those are free are implemented with the VM for the job request
   b. The comparison of the performance in the available VMs the cost should be minimal in the VM
   c. The reduction of the time caused by the reduced energy consumption.

Through the consideration of the job the VM is assigned in the network.

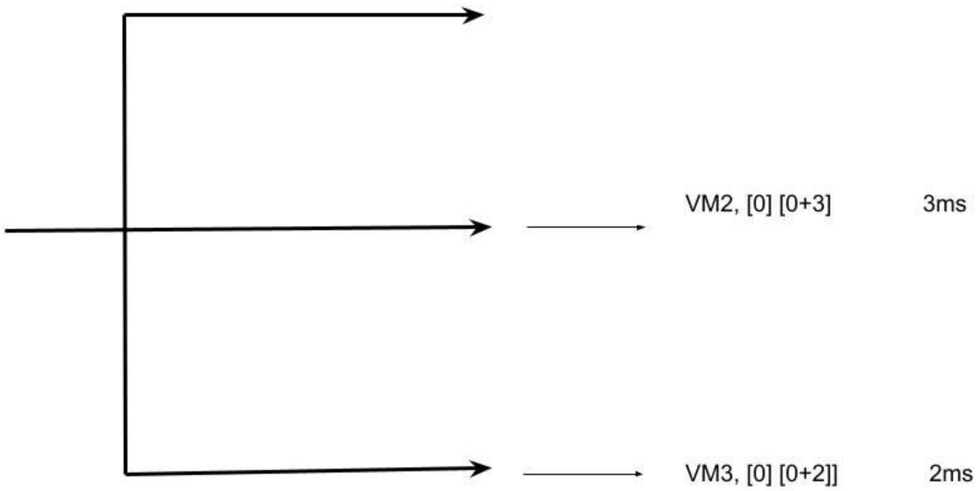

Figure 2. VM assignment

In figure 3 presented about the assigned jobs for the VM in the proposed MFCLS. The estimation is based on the computation of the tasks and jobs in the network.

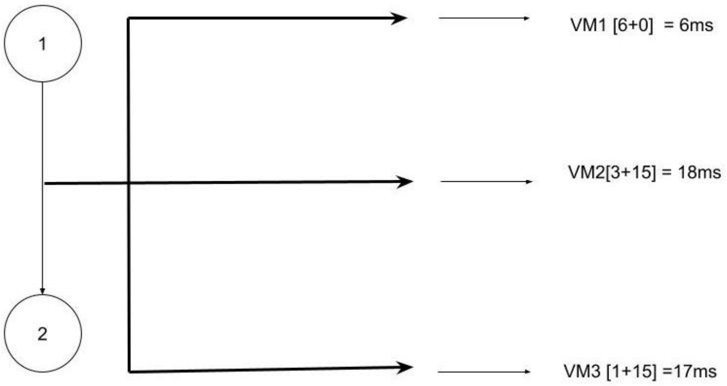

Figure 3. Assignment of Job in VMs

The medical data is more sensitive and stable with the selection of features in the created medical datasets. In order to make this approach stable, we need to pay attention on the probability measures of the attributes over the small subsets of data to understand the contribution of the attributes. Based on the consideration of the scenario the entropy based technique need to be implemented for the processing of large amount of data. In figure 4 data analytics model implemented for the proposed MFCLS model is presented.

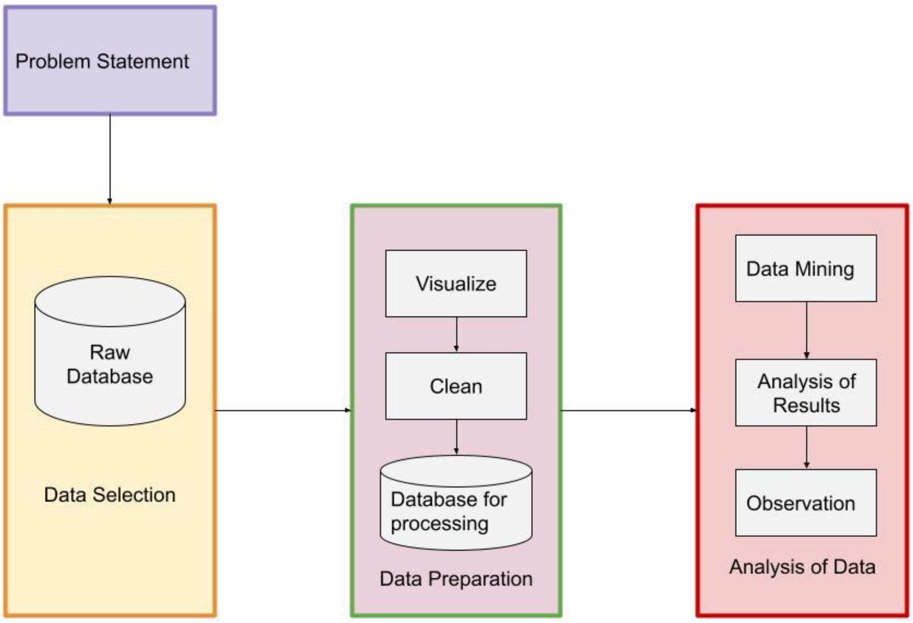

Figure 4. Process in MFCLS data analytics model

In figure 5 the communication interface for the proposed MFCLS model is presented. The developed architecture model comprises of the information exchange between patient and doctor towards the decision making process.

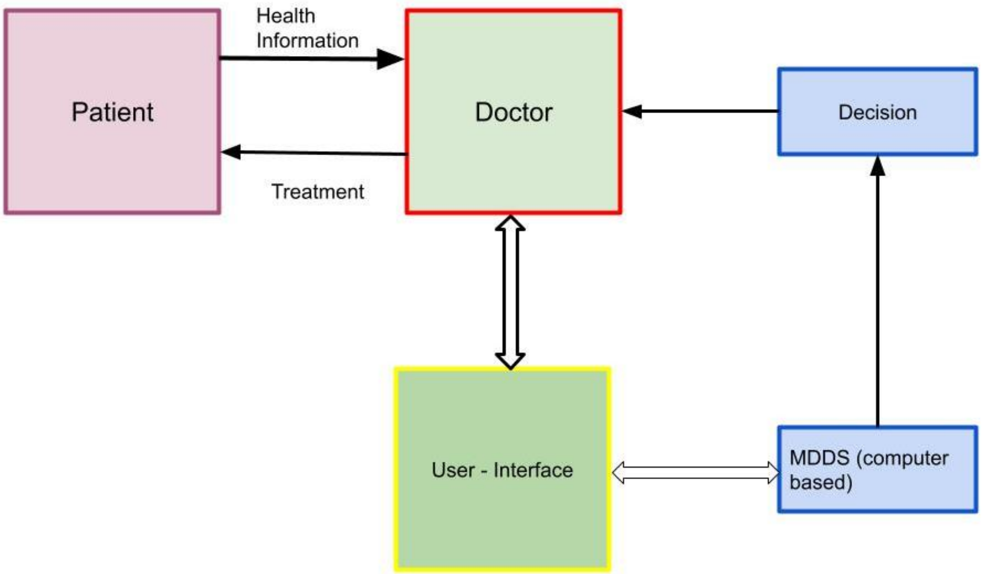

Figure 5. Flow of fog architecture in the MFCLS

In the table 1 the assigned number of job cost for the VM is presented. The assigned number of job is considered as 10 for the VM count of 3.

Table 1. Job cost for different VM

Job No	VM job cost		
	VM1	VM2	VM3
1	1	3	7
2	3	4	1
3	2	3	2
4	1	2	3
5	4	2	5
6	3	1	4
7	11	10	11
8	4	7	8
9	10	3	7
10	7	2	5

The calculation is explained using the following architecture shown in Figure 4.1 and Table 1 shows the costing of different VMs

The total cost of execution is:

$$Total\ cost = E_c + T_c \tag{8}$$

Where:

$$E_c = Execution\ cost$$
$$T_c = Transfer\ cost$$

3.1 Entropy based feature selection

The data cleaning operations like handling missing values in the attribute of the datasets, removing redundant features/irrelevant features from the datasets, handling data consistency, features dimension reduction etc. are performed with the proposed MFCLS evaluate the healthcare dataset (D). With the proposed MFCLS scheme feature selection is performed with the discretise of the by Minimum-information loss (MIL) for the 3 different subsets D1, D2 and D3 with data partitioning in the equip class data distribution. The feature selection in the entropy based model for the estimation of the features is presented in figure 6.

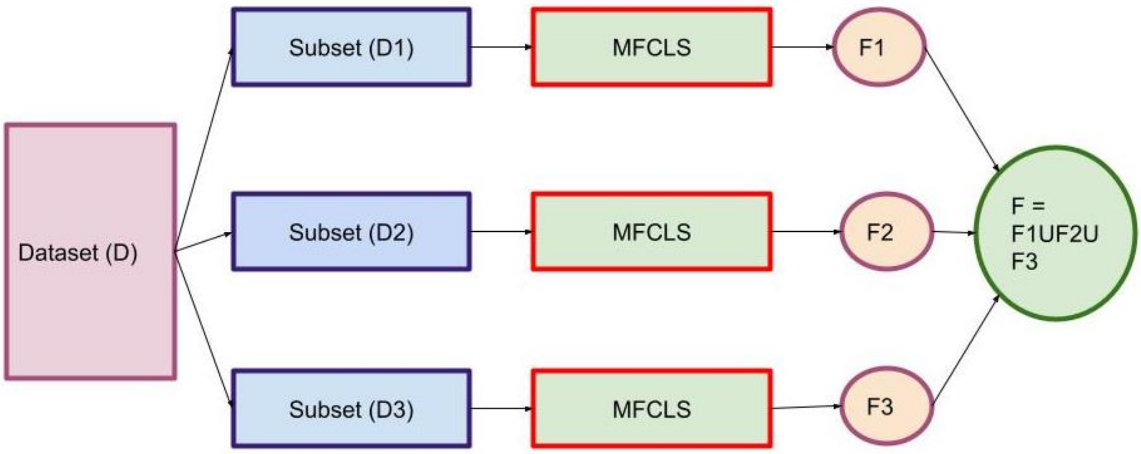

Figure 6. Feature Selection with MFCLS

Equi-class Distribution

The proposed MFCLS model integrates the equi-class distribution for the healthcare medical dataset evaluation with the training set of data.

In the feature selection process the proposed MFCLS model perform the classification presented as follows:

1. D1 (1st sub-set): In this dataset, the 30% of data were randomly derived from the complete healthcare data D with the elimination of the selected instances D.
2. D2 (2nd sub-set): In the random manner 30% of data were utilized for the selected data in D with the distinct instances D2 in the dataset D.
3. D3 (3rd sub-set): The remaining data in the dataset D aprat from D1 and D2 are placed in the D3. This dataset is completely independent form for analysis.

In the proposed MFCLS model the dataset (D) comprises of the different attributes A_i, $i = 1, \ldots., n$. Consider the feature set denoted as $F_s = \{A_1, A_2, \ldots \ldots, A_n\}$ for the dataset D those are classified in the 3 subsets D_1, D_2 *and* D_3 for the three different features in the distinct data subsets. Consider the feature set represented as F_1, F_2 *and* F_3 with the value $F_k(k = 1,2,3) = F_s$. In the obtained set the resulted features F denoted as F_1, F_2 *and* F_3.

Consider the classification problem P with the total attributes count of n stated as $A_{i(i=1,\ldots\ldots,n)}$. The feature set is represented as F with the value of $F = F_s = \{A_1, A_2, \ldots \ldots, A_n\}$.

1. for every subset data $(D_k, k = 1,2,3)$ do
2. begin
3. for each attributes in the data $A_{i(i=1,\ldots\ldots,n)}$ do
4. begin
5. Compute the information gain the in the healthcare data A_i (ie $Gain(S, A_i)$ as presented in equation (9)

$$Gain\ (S, A_i) = Entropy\ (S) - \sum_{v_j \epsilon A_i} \frac{|S_{v_j}|}{|S|} Entropy\ (S_{v_i}) \qquad (9)$$

where, $v_j(j = 1, \ldots.k)$ represented as the attribute Ai with the computation of entropy $Entropy\ (S) = \sum_{m=1}^{c} p_m\, log_2 P_m$ where number of exapmles P is computed in the sample P. The non-zero probability is denoted as S_m with the class m, out of c.

end for

6. Compute $r = ((\max_{Gain}(S, A_i) - \min\ \_Gain(S, A_i)ln,\ i = 1, \ldots\ldots, n$
7. For every healthcare data attributes $A_{i(i=1,\ldots\ldots,n)}$ do

Begin

8. If $Gain\ (S, A_i)_{(i=1,....,n)} < r$, then feature set need to be updated (F_k)as: $F_k = F_k^- \{A_i\}$ // discard the attributes A_i from F

end for

end for

9. $F = F_1 U F_2 U F_3$

/*F comprises of the non-common and common attributes of the healthcare data F_1, F_2 and F_3. The balanced dataset comprises of the subsets to F_1, F_2 and F_3 are those are balances with the common attributes F_1, F_2 and F_3 increases. */. The assigned jobs for the propised MFCLS is presented.

Algorithm 1: Job Assignment with MFCLS

Input: List of Host (HL), List of Job (JL) and Deadline
JL.Decreasing with sorting_Optimization ()
for
Every job in JL, update list do
Minimum Power_Utilization ← Maximal power allocated
Assigned to host <--null
Sufficient job resources available
Estimate Power ← Power as (H,J)
if
Power < utilized power_then
Compute fitness (HL, Job, Deadline);
if fitness function is satisfied as fitness_Fxn
Optimize Host value
end if
end if
Return;
Optimize the HL & JL values
end
end for

With the proposed MFCLS laod scheduling with the optimization model for the job assignemtn is presented.

Algorithm 2: Optimization for load scheduling
Input: Host List Optimization (HL) and Job list Optimization (JL) Configure the ANN parameters such as Neurons (N): Epochs (E) Performance Estimation with validation; Gradient; MSE; Mutation Training of sequences with Levenberg Marquardt as (Trainlm) Random division of data for each job compute host Based on host ability the Group (G) compute based on categories end Initialize the data group and training using ANN Set the parameters for training Net = train (Groups, Jobs) Return; Train ANN for the allocation of job End

4. Results and Discussion

The proposed MFCLS model comprises of the medical data processing based on the assigned job in the resource scheduling. The parameters concentrated on the privacy improvement in the healthcare data.

Security Lifecycle Review (SLR): In the every VM system the task completion time is accomplished based on the job model. The SLR in the calculated using the equation (10)

$$SLR = \sum_{k=0}^{n} Job\ execution\ time \qquad (10)$$

Where,

n = number of jobs in the mode

SLVMR: The job is executed by the one VM at the time it provides the ration of the SLR to the computed time as it is given in equation (11) as follows:

$$SLVMR = \frac{SLR}{Job\ time\_one\ VM} \qquad (11)$$

EC: The total energy consumed to execute the all tasks in the sequences.

ECR: The tasks are executed by the one VM at a time and its is computed using the formula (12)

$$ECR = \frac{EC}{Energy\ Consumed} \qquad (12)$$

Table 2. Health care data for scheduling

Medical Data	**Attributes Number**	**Class Number**	**Sample Number**
Breast Cancer	9	2	698
Dermatology	32	6	367
Ecoli	7	7	334
Heart	12	5	293
Heart (Swiss)	12	4	121
Heart (Cleveland)	12	4	301
Hepatitis	18	2	153
Liver Disorder	6	1	343
Lung Cancer	55	3	31
Lymphography	17	4	147
Thyroid	5	2	213
Diabetes	8	1	754
Primary Tumor	16	21	337
Sick	27	2	3772

In the table 2 different datasets such as dermatology, breast cancer, sick and primary tumor, heart, lung cancer and Hepatitis based on consideration of different datasets dimensions. The feature selection is implemented and evaluated with the 14 datasets. The presented datset feature selected are presented in table 3. The dataset for analysis is presented in the dataset defined as DN, number of instances (NI), feature number (NF), selected features number (NSF) and selected features (SF).

Table 3. Feature Selection for medical data

Medical Data	**Number of instances**	**Number of features in the original datasets**	**Number of Selected features**	**Selected Features**
Breast Cancer	698	9	6	1,3,4,5,6,9
Dermatology	367	33	21	2,3,4,5,6,7,9,13,14,15,16,19,20,21,22,26,27,28,29,30,33
Ecoli	334	8	6	1,2,3,5,6,7
Heart	293	12	6	2,3,6,9,10,11
Heart (Swiss)	121	12	3	9,10,11
Heart (Cleveland)	301	12	7	2,3,8,9,10,11,12,13
Hepatitis	153	19	8	1,2,6,11,12,14,17,18,19

Liver Disorder	343	6	2	3,5
Lung Cancer	31	57	18	1,3,5,9,13,14,15,20,21,25,26,38,41,45,48,50,56
Lymphography	147	17	10	1,2,7,8,9,11,13,15,16,18
Thyroid	213	5	5	1,2,3,4,5
Diabetes	754	7	4	2,6,7,8
Primary Tumor	337	16	11	1,2,3,4,5,7,9,10,13,15,16,17
Sick	3772	28	8	1,8,10,14,15,20,24,29

For more clarity about the feature reduction work performed in this study, the important terminologies that are used here are presented below.

• Original Features: - Features that are available in the original datasets (uploaded in the UCI repository) are known as original Features. These are, indeed, decided by the experts (physicians) to diagnose the diseases.

Selected Features: - Features selected from each original dataset by removing noisy (irrelevant/redundant) features using suitable feature reduction approach, are known as selected features.

Independent Features: - Features that are not dependent among themselves, are known as independent features. Each such features are dependent with target feature, i.e., their contribution in decision making is high enough.

Let us consider the disease dataset named Heart (Swiss) dataset for explanation. In UCI repository, 13 non-target attributes are decided as the processed features by the experts (out of 76 unprocessed features). These 13 features are now treated in the present experiment as the original features. These are renamed here as A1, A2, A3, …, A13 (in short, 1, 2, 3, …, 13) with description as follows.

1. Age (A1): Age of patient in years.
2. Sex (A2): 0 = male; 1 = female).
3. Tc (A3): Type of chest pain under 4 types

1 – Angina typical
2 – Angina atypical
3 – Non-anginal pain
4 - Asymptomatic

4. Bp (A4): Blood pressure measured with mmHg in the hospita;
5. Cholestral (A5): Serum cholesterol measured in mg/dl
6. Fbs(A6): Blood sugar in fasting if 120mg/dl

1 – True
0 - False

7. ECresults (A7): Electrocardiographic results in the resting states

0 – Normal
1 – ST – T abnormality in wave >0.05mV
2 – Estes Criteria for the left ventricular hypertrophy

8. Maxrate (A8): Maximal heart rate obtained
9. Angex (A9) – Angina induced by angina

0 – No
1 – Yes

10. Depeak (A10) – The induced depression for the rest relative exercises
11. Slope (A11) – The exercise peak in the segmented T-wave

1 – upslope
2 – Flat
3 – Downslope

12. Cofl(A12) – Fluoroscopys major vessels ranges from 0 – 3
13. Blodis (A13) – Thalassemia

1 – Normal
2 – defect fixed

3 – Defect those are reversible
Thus, the selected features by the adopted hybrid feature selection approach are namely, 9 (A9), 10 (A10) and 11 (A11).
The following are used for the comparison of the effective hybrid feature selection approach
1. Three states of the learners used – C4.5, RIPPER, naïve Bayes
2. The cross-validation in 10-fold manner for the 10 runs with the feature selection with the classification accuracy with the desire standard deviation.
3. Number reduced features by the adopted approach

Table 4. Comparison of Attributes

Medical Data	Number of attributes	J48	JRIP	Proposed MFCLS
Breast Cancer	9	73.28± 6.03	75.67± 4.74	88.74± 1.86
Dermatology	32	90.13± 3.33	88.74± 2.97	90.35± 5.83
Ecoli	7	81.32± 5.69	83.76± 3.78	91.56± 2.78
Heart	12	77.23± 7.85	82.73± 5.74	88.93± 6.83
Heart (Swiss)	12	34.56± 12.75	69.56± 9.06	78.45± 7.34
Heart (Cleveland)	12	75.67± 6.46	81.46± 4.62	88.93± 3.12
Hepatitis	18	76.32± 6.27	77.62± 3.02	92.58± 0.91
Liver Disorder	6	63.49± 7.39	67.83± 5.73	89.32± 1.36
Lung Cancer	55	68.35± 22.63	71.56± 8.73	93.46± 2.78
Lymphography	17	74.83± 11.35	73.67± 4.67	86.39± 1.93
Thyroid	5	86.57± 5.62	89.74± 9.76	91.36± 4.67
Diabetes	8	74.57± 4.78	81.49± 8.19	88.64± 6.83
Primary Tumor	16	43.68± 6.84	73.78± 3.67	93.67± 1.04
Sick	27	88.94± 1.34	89.56± 4.03	93.16± 2.74

The number of job assigned for the VM in the fog architecture is presented in figure 7. Based on the assigned VM in the 1, 2 and 3 the computation cost is estimated.

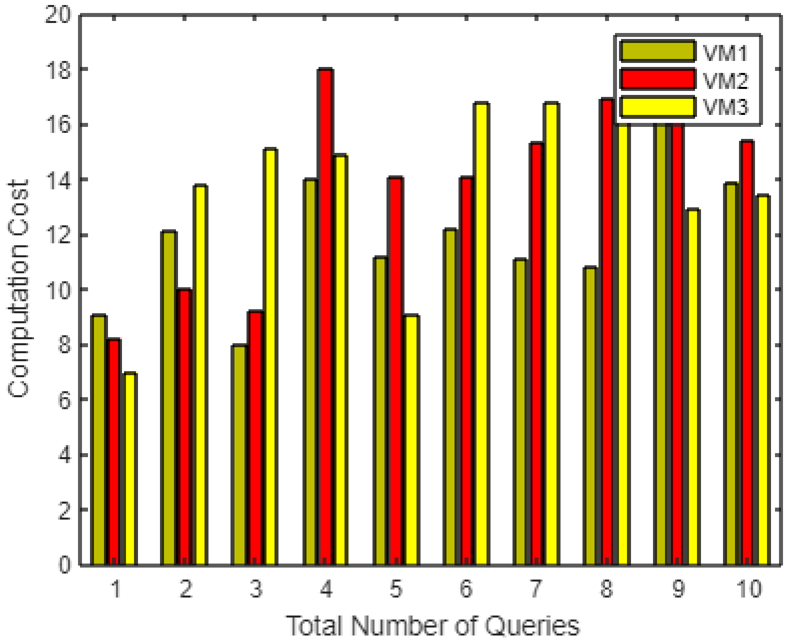

Figure 7. Computation Cost for different VM

The data privacy is estimated comparatively with the proposed MFCLS model are presented in figure 8.

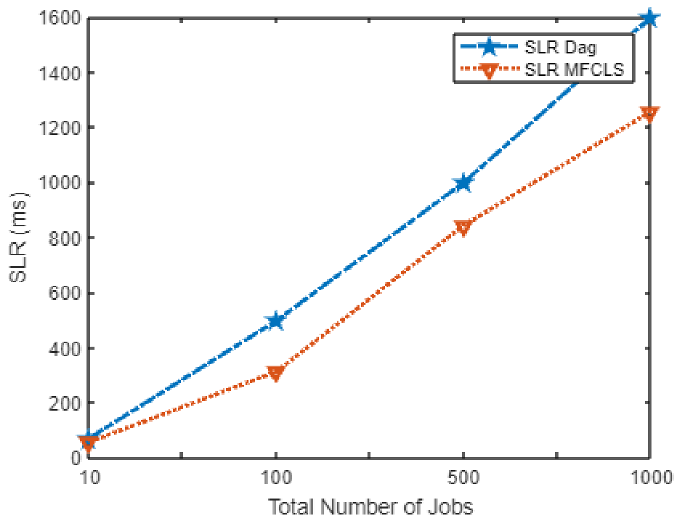

Figure 8. Comparison of SLR

The privacy estimation of the SLR expressed that the proposed MFCLS model achieves the higher privacy compared with the DAG model. In figure 9 and figure 10 the comparison of the energy consumption for the total number of jobs and each job is presented.

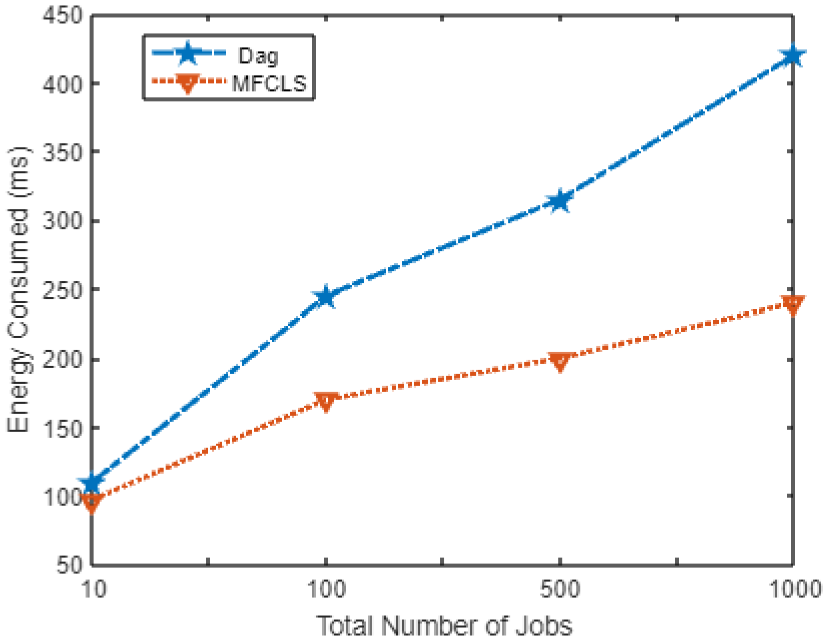

Figure 9. Energy Consumption for total jobs

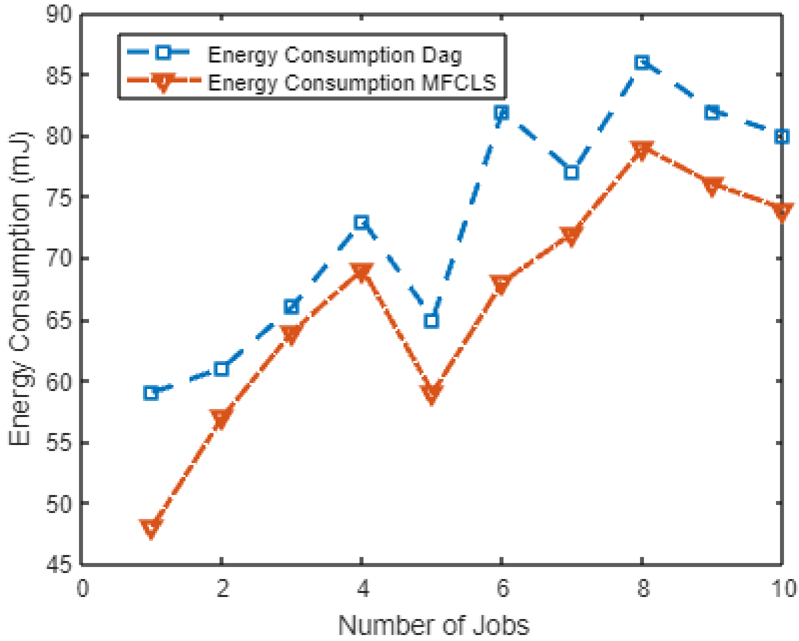

Figure 10. Energy Consumption for each individual jobs

As the MFCLS model involved in the resource scheduling for the assigned VM the jobs assigned for the each individual is computed as presented in the figure 11 and figure 12.

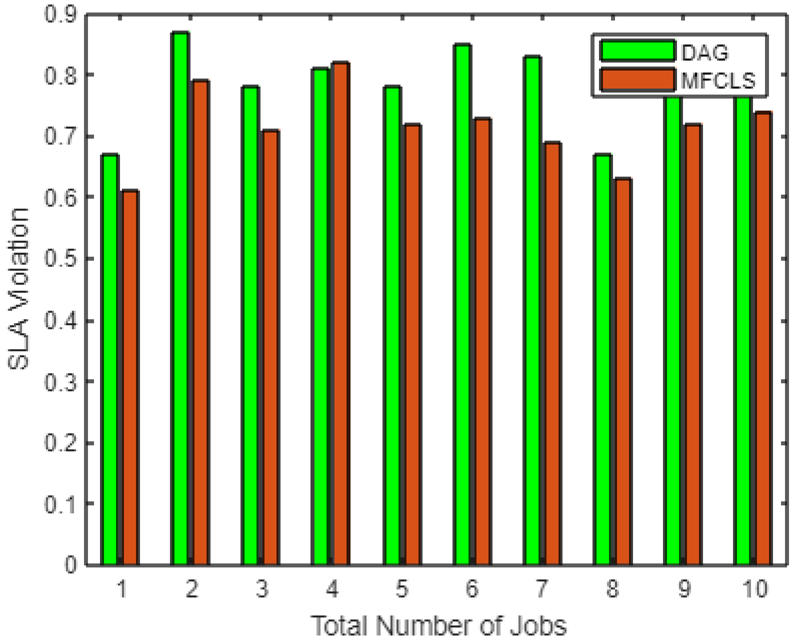

Figure 11. SLA violation

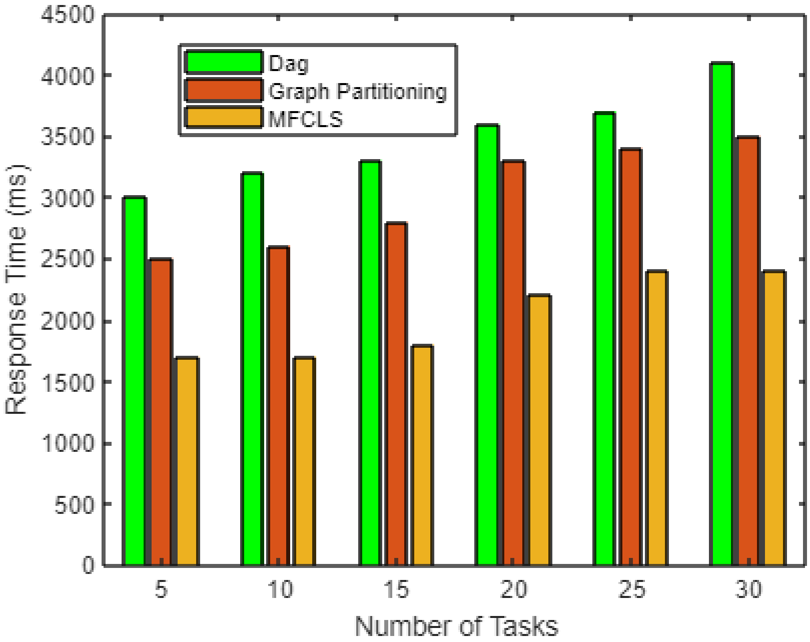

Figure 12. Comparison of Response Time

The SLA violation comparison expressed that the proposed MFCLS model exhibits the minimal SLA violation compared with DAG. The accuracy for the selected dataset is examined with the consideration of the learners rate. The computation accuracy increases with the elimination of the features those are redundant. Particularly, with the implementation of the naïve bayes classifier the cases performance is increases significantly with the feature filtration technique. The naïve bayes classifier exhibits the significant performance for the datasets with consideration of independent features with the selection of feature attributes. Actually, the estimation of the features significantly eliminates the set of filteration approach with the reduced information gain. In other words, the demand is based on the consideration of the dependent features with the elimination of the information in the datasets. The defined dataset need to be more reliable with effective standard deviation value for the yielded accuracy level in the classifier compared with the standard deviation value considered for the original ones. With the feature reductio in the proposed MFCLS model learning rate need to be minimized with the higher extent for the induced small rule. In some datasets, improvement in the used metrics yield by some learners (not for all the chosen learners) are observed unchanged or very less or acceptably down, but amount (%age) of dimension reduction (i.e., noise reduction) is considerably good. This may be due to the adopted learning strategies by the learners that usually desire more features while training.

5. Conclusion

Fog computing technology exhibits the advancement over the CC implemented with the effective 5G communication in the healthcare applications. In the healthcare application fog computing technology concentrated on the reduced energy consumption with the load optimization. The proposed MFCLS model comprises of the optimization based model for the healthcare data processing. The proposed MFCLS model uses the entropy based feature selection model for the optimization of the ABC algorithm with the supervision and stabilization of the load in the fog computing network with the reduced energy consumption and task runtime length. The developed MFCLS model stabilize the load in the fog computing with the improved performance in the framework with the recurring framework of 1000 jobs. The formulated fitness function exhibits the significant performance towards minimization of the task load in the VM with the initiation and execution time. The algorithm used for the ranking comprises of the connected components for the total energy consumed for the completion of the tasks. The proposed MFCLS algorithm uses the schedule length estimation for the developed MFCLS architecture model to increases the data privacy in the fog. The proposed MFCLS model exhibits the smallest time frame compared with the existing algorithm. Additionally, the analysis of the proposed MFCLS model comparatively examined with the DAG model. The performance analysis expressed that the proposed MFCLS model achieves the SLRVM for the data achieves the value of 28 in case of the conventional DAG model it is achieved as 88. Additionally, ECRVM model achieves the task value of 43 wheareas the DAG model provides 7. The comparative examination confirmed that proposed MFCLS model exhibits ~6% increased performance than the conventional DAG technique with the adequate data privacy in the heathcare data implemented in fog architecture.

References

[1] Sivasangari, A., Lakshmanan, L., Ajitha, P., Deepa, D., & Jabez, J. (2021). Big Data Analytics for 5G-Enabled IoT Healthcare. In Blockchain for 5G-Enabled IoT (pp. 261-275). Springer, Cham.

[2] Kamruzzaman, M. M., Yan, B., Sarker, M. N. I., Alruwaili, O., Wu, M., & Alrashdi, I. (2022). Blockchain and Fog Computing in IoT-Driven Healthcare Services for Smart Cities. Journal of Healthcare Engineering, 2022.

[3] Chu, K. M. Understanding Ioht and Edge/Fog Computing Solutions for Smart In-Home Remote Healthcare. Fog Computing Solutions for Smart In-Home Remote Healthcare.

[4] Tselios, C., Politis, I., Amaxilatis, D., Akrivopoulos, O., Chatzigiannakis, I., Panagiotakis, S., & Markakis, E. K. (2022). Melding Fog Computing and IoT for Deploying Secure, Response-Capable Healthcare Services in 5G and Beyond. Sensors, 22(9), 3375.

[5] Prasad, V. K., Tanwar, S., & Bhavsar, M. D. (2021). Advance cloud data analytics for 5G enabled IoT. In Blockchain for 5G-Enabled IoT (pp. 159-180). Springer, Cham.

[6] Muneeb, M., Ko, K. M., & Park, Y. H. (2021). A fog computing architecture with multi-layer for computing-intensive IoT applications. Applied Sciences, 11(24), 11585.

[7] Singh, J., Singh, P., & Gill, S. S. (2021). Fog computing: A taxonomy, systematic review, current trends and research challenges. Journal of Parallel and Distributed Computing, 157, 56-85.

[8] Sicari, S., Rizzardi, A., & Coen-Porisini, A. (2022). Insights into security and privacy towards fog computing evolution. Computers & Security, 102822.

[9] Sodhro, A. H., & Zahid, N. (2021). AI-enabled framework for fog computing driven e-healthcare applications. Sensors, 21(23), 8039.

[10] Sarrab, M., & Alshohoumi, F. (2022). Assisted Fog Computing Approach for Data Privacy Preservation in IoT-Based Healthcare. In Security and Privacy Preserving for IoT and 5G Networks (pp. 191-201). Springer, Cham.

[11] Alzoubi, Y. I., Al-Ahmad, A., & Kahtan, H. (2022). Blockchain technology as a Fog computing security and privacy solution: An overview. Computer Communications, 182, 129-152.

[12] Ahanger, T. A., Tariq, U., Ibrahim, A., Ullah, I., Bouteraa, Y., & Gebali, F. (2022). Securing IoT-Empowered Fog Computing Systems: Machine Learning Perspective. Mathematics, 10(8), 1298.

[13] Alwakeel, A. M. (2021). An overview of fog computing and edge computing security and privacy issues. Sensors, 21(24), 8226.

[14] Samann, F. E., Abdulazeez, A. M., & Askar, S. (2021). Fog Computing Based on Machine Learning: A Review. International Journal of Interactive Mobile Technologies, 15(12).

[15] Patwary, A. A. N., Naha, R. K., Garg, S., Battula, S. K., Patwary, M. A. K., Aghasian, E., ... & Gong, M. (2021). Towards secure fog computing: A survey on trust management, privacy, authentication, threats and access control. Electronics, 10(10), 1171.

[16] Laghari, A. A., Jumani, A. K., & Laghari, R. A. (2021). Review and state of art of fog computing. Archives of Computational Methods in Engineering, 28(5), 3631-3643.

[17] Abdulqadir, H. R., Zeebaree, S. R., Shukur, H. M., Sadeeq, M. M., Salim, B. W., Salih, A. A., & Kak, S. F. (2021). A study of moving from cloud computing to fog computing. Qubahan Academic Journal, 1(2), 60-70.

[18] Tewari, N., & Budhani, S. K. (2022). Analysis of Risk and Security Within Fog Computing-Enabled e-Healthcare System in Uttarakhand. In Cyber Security and Digital Forensics (pp. 21-30). Springer, Singapore.
[19] Veni, T. (2022). Quantum-Based Resource Management Approaches in Fog Computing Environments: A Comprehensive Review. Mobile Computing and Sustainable Informatics, 743-752.
[20] Ngabo, D., Wang, D., Iwendi, C., Anajemba, J. H., Ajao, L. A., & Biamba, C. (2021). Blockchain-based security mechanism for the medical data at fog computing architecture of internet of things. Electronics, 10(17), 2110.
[21] Hewa, T., Braeken, A., Liyanage, M., & Ylianttila, M. (2022). Fog computing and blockchain-based security service architecture for 5G industrial IoT-enabled cloud manufacturing. IEEE Transactions on Industrial Informatics, 18(10), 7174-7185.
[22] Hewa, T., Braeken, A., Liyanage, M., & Ylianttila, M. (2022). Fog computing and blockchain-based security service architecture for 5G industrial IoT-enabled cloud manufacturing. IEEE Transactions on Industrial Informatics, 18(10), 7174-7185.
[23] Karakaya, A., & Akleylek, S. (2021). A novel IoT-based health and tactical analysis model with fog computing. PeerJ Computer Science, 7, e342.