Aerospace Systems (2022) 5:615–624
https://doi.org/10.1007/s42401-022-00158-6

ORIGINAL PAPER

Advantage of IRNSS S-band signal for GBAS applications in adverse ionospheric storm conditions

Supraja R. Ammana[1] · M. R. Sujimol[2] · Komal Kumar Songala[1] · A. D. Sarma[1]

Received: 3 April 2022 / Revised: 27 July 2022 / Accepted: 31 July 2022 / Published online: 20 August 2022

Abstract
One of the significant features that makes Indian Regional Navigation Satellite System (IRNSS) distinct from most of the other satellite navigation systems is the use of S band signal along with the L-band signals. The fact that S-band signals experience lesser ionospheric delays and therefore lesser ionospheric gradients than L-band signals is of great advantage, more-so in ionospheric storm conditions. The advantage of using IRNSS S-band signal under ionospheric storm conditions for Ground Based Augmentation System (GBAS) applications is investigated in this paper. The analysis is carried out by computing a GBAS parameter namely sigma-vertical-ionospheric-gradient (σ_{vig}) using data from low-latitude stations in Hyderabad, India (17.3850° N, 78.4867° E) under both quiet and storm ionospheric conditions and then estimating vertical protection levels (VPLs). It is observed that even in storm conditions, σ_{vig} is only 5.08 mm/km with S1 as against 22.80 mm/km due to L5. Also, the VPLs are 0.65 m less than those with L5 and are well within the alert limits. This work carries significance in view of GBAS being planned at several low-latitude airports.

Keywords IRNSS · S-band · GBAS · Ionosphere

1 Introduction

Indian Space Research Organization (ISRO) developed IRNSS with the main aim of providing Restricted Service (RS) and Standard Positioning Service (SPS) for authorized users and civilian users, respectively. Now IRNSS is popularly called with its operational name as Navigation with Indian Constellation (NavIC). The NavIC satellites transmit SPS signals on S-band (2492.028 MHz) and L5 band (1176.45 MHz) [1]. The ionospheric delay is proportional to the number of free electrons along the propagation path, which is known as the total electron content (TEC). Further, as the ionosphere is a dispersive medium, the delay is frequency dependent and is inversely proportional to square of the frequency and therefore S-band signals have got certain obvious benefits over L-band, especially in the context of ionosphere induced errors. The advantages include smaller scintillations, lower ionospheric delays and therefore less ionospheric gradients in S-band than L-band [2]. The positive aspect of using S-band signal in the context of Ground Based Augmentation System (GBAS) is explored in this paper.

GBAS is being implemented worldwide in airports for precision approach (PA) and landing of aircrafts. Integral component of GBAS architecture is the GBAS ground facility (GF), consisting of GNSS receivers, installed at appropriate locations in the airport and VHF data broadcast (VDB) transmitter [3, 4]. GNSS satellites provide ranging signals to the reference receivers and airborne users. The pseudorange measurements obtained by these receivers are processed in GF to generate differential corrections corresponding to each satellite signal. These differential corrections are applied by the airborne users to the pseudoranges to obtain an accurate position estimate. The differential corrections will be able to completely eliminate the ionospheric error only when the ionospheric delays at the user and the GF are similar. Difference of ionospheric delay between the user and GF (ionospheric spatial gradients), gives rise to residual ionospheric error at the user. The bounds on the residual ionospheric error are assessed by the airborne user with the help of a bounding sigma, namely, sigma-vertical-ionospheric-gradient (σ_{vig})). This parameter is broadcast by

✉ A. D. Sarma
ad_sarma@yahoo.com

1 Chaitanya Bharathi Institute of Technology (CBIT), Hyderabad, India

2 Delhi Earth Station, Space Applications Centre (SAC), ISRO, New Delhi, India

Transactions on Emerging Telecommunications Technologies / Volume 34, Issue 1 / e4680

RESEARCH ARTICLE

Artificial intelligence framework-based ultra-lightweight communication protocol for prediction of attacks in Internet of Things environment

Mounika Jammula, Venkata Mani Vakamulla, Sai Krishna Kondoju

First published: 15 November 2022
https://doi.org/10.1002/ett.4680

Abstract

The Internet of Things is an enhanced intelligent infrastructure that is created through the use of a number of different devices that are capable of self-organization. Data privacy and protection are the primary problems brought up by this enormous network. In addition, the devices in the network with constrained resources are battery-powered, thus they only have a limited capacity to store resources internally. Because of this, it is vital to discover solutions that are more resource-optimized and connected to security in order to solve the issues that were caused by the network. The devices are being forced to maintain more complicated cryptographic algorithms for higher security. Therefore, this article proposes an integrated communication protocol that requires only symmetric key-based cryptography with a deep learning convolutional neural network, which is employed for predicting the normal and attacked data from the input requested data. Here, the logical map is used to produce the symmetric keys due to its resistance against the attacks of key reset and device capture. This integrated model provides an ultra-lightweight communication protocol with reduced attack detection time while enhancing the attack detection parameters such as precision, accuracy, F1-score, and recall.

CONFLICT OF INTEREST

Authors declared that there is no conflict of interest.

Open Research

Sensors and Actuators A: Physical
Volume 353, 16 April 2023, 114232

Electrochemical modified Pt nanoflower @rGO for non- enzymatic electrochemical sensing of glucose

Saravanan Gengan [a], R.M. Gnanamuthu [b], Sanjay Sankaranarayanan [c], Venumbaka Maneesh Reddy [d], Bhanu Chandra Marepally [e], Ravi Kumar Biroju [f] [g]

Show more

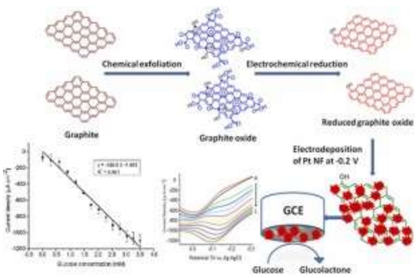
Share Cite

https://doi.org/10.1016/j.sna.2023.114232
Get rights and content

Abstract

Since lower danger of biorecognition element degradation, enzymes-less glucose have the potential for more reliable in vivo activity, but it suffers due to lack of linear response and poor selectivity. We made attempt to improve selectivity, linear response and stability, environmentally benign electrochemical method adopted to fabricate Pt nanoflowers (PtNF) anchored on rGO modified GCE (PtNF-rGO/GCE). The PtNF-rGO/GCE electrode demonstrated good glucose electrooxidation in alkaline solution, with a linear range, sensitivity and detection limit are up to 3.5 mM, 335.5 μA mM^{-1} cm^{-1} and 53 μM (S/N=3) respectively. The PtNF-rGO/GCE electrode is not only selective also inhibit interfering molecules like uric, dopamine, ascorbic acid. This allows for broadly sensitive, work at low-potential, stable, and quick glucose current detection, which is capable for the expansion of non-enzymatic glucose detectors.

Graphical Abstract



Download : Download high-res image (172KB)
Download : Download full-size image

Introduction

Research on novel high-performance devices has increased for medical diagnostics, environmental control applications [1], [2]. Electrochemical sensors provide a number of benefits over traditional analytic methods, including cheap cost, the capacity to detect substances in real time, simplicity, and the ability toward miniaturize. Due to their unique physicochemical properties, researchers have combined the benefits of electrochemical artifice with nanomaterials [3], [4]. The overall kinetics of electrode reaction at bare Pt surface is too sluggish to produce significant faradaic currents, in addition to the low sensitivity, strongly impaired by adsorbed chloride ions and chemisorbed intermediates and quickly block the electroactive surface. Pt nanomaterials have been used as electrochemical sensing elements because of their good biocompatibility, enhance electron transfer rate, excellent electrical and thermal conductivity, excellent chemical stability, and large specific surface area. Platinum nanoparticles have unique properties such as their surface effect, volume effect, quantum size effect, and macroscopic quantum tunneling effect [5], [6], [7]. Materials in nanosize improve electrochemical sensitivity, particularly of kinetically regulated electrochemical processes [8], [9]. Park et al. [10] reported on Pt nanostructures boost electrode sensitivity to the point where enzyme-less glucose sensing with kinetically regulated electrooxidation is possible. Choices of approaches for obtaining Pt nanoparticles have been available in the literature, using surfactant syntheses, template assisted dealloying

ACS OMEGA

http://pubs.acs.org/journal/acsodf

Article

Defective Graphene/Plasmonic Nanoparticle Hybrids for Surface-Enhanced Raman Scattering Sensors

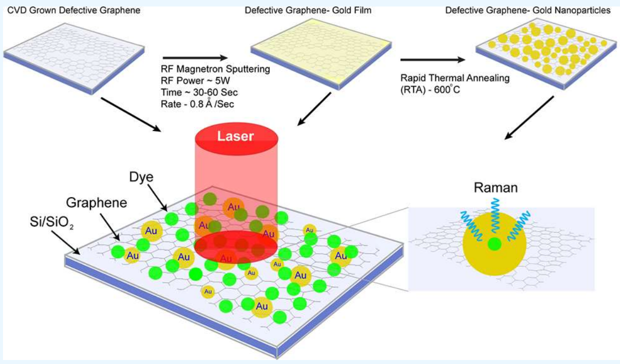
Ravi K. Biroju,* Bhanu Chandra Marepally, Pariksha Malik, Soumen Dhara, Saravanan Gengan, Dipak Maity, Tharangattu N. Narayanan, and Pravat K. Giri*

Cite This: *ACS Omega* 2023, 8, 4344−4356

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Two-dimensional−zero-dimensional plasmonic hybrids involving defective graphene and transition metals (DGR-TM) have drawn significant interest due to their near-field plasmonic effects in the wide range of the UV−vis−NIR spectrum. In the present work, we carried out extensive investigations on resonance Raman spectroscopy (RRS) and localized surface plasmon resonance (LSPR) from the various DGR-TM hybrids (Au, Ag, and Cu) using micro-Raman, spatial Raman mapping analysis, high-resolution transmission electron microscopy (HRTEM), and LSPR absorption measurements on defective CVD graphene layers. Further, electric field (E) mappings of samples were calculated using the finite domain time difference (FDTD) method to support the experimental findings. The spatial distribution of various in-plane and edge defects and defect-mediated interaction of plasmonic nanoparticles (NPs) with graphene were investigated on the basis of the RRS and LSPR and correlated with the quantitative analysis from HRTEM, excitation wavelength-dependent micro-Raman, and E-field enhancement features of defective graphene and defective graphene-Au hybrids before and after rapid thermal annealing (RTA). Excitation wavelength-dependent surface-enhanced Raman scattering (SERS) and LSPR-induced broadband absorption from DGR-Au plasmonic hybrids reveal the electron and phonon interaction on the graphene surface, which leads to the charge transfer from TM NPs to graphene. This is believed to be responsible for the reduction in the SERS signal, which was observed from the wavelength-dependent Raman spectroscopy/mappings. We implemented defective graphene and DGR-Au plasmonic hybrids as efficient SERS sensors to detect the Fluorescein and Rhodamine 6G molecules with a detection limit down to 10^{-9} M. Defective graphene and Au plasmonic hybrids showed an impressive Raman enhancement in the order of 10^8, which is significant for its practical application.



1. INTRODUCTION

The construction of transition metal (TM) plasmonic nanostructures (PNSs) incorporated with graphene is attracting considerable research attention for enhanced photocatalysis and photovoltaic applications.[1−3] However, ultrathin films and nanoparticles (NPs) of TM interfaced with graphene have been exploited for interesting properties in the graphene research community.[4,5] Various chemical and physical approaches have been developed to design graphene-plasmonic hybrids for enhanced optical and optoelectronic characteristics. Most of the experiments on fabrication were carried out by chemical approaches, and the contribution of graphene and its structural evolution is lacking in the recent reports on the graphene-TM PNSs.[5,6] Understanding the interaction of various TM films and NPs over the graphene layer preserving its high structural quality and the role of intrinsic defects using resonance Raman scattering (RRS) and localized surface plasmon resonance (LSPR) absorption techniques is challenging. On the other hand, the uniqueness of graphene having a high charge carrier mobility and a wide range of optical transparency over the visible to near-infrared (NIR) region makes it suitable for the resonance of surface plasmon originating from the TM PNSs in the visible excitation energy. These graphene-TM hybrid functional materials can show enhanced visible light photocatalytic activity and efficient photovoltaic effects incorporating various photosensitive nanomaterials.[7] As a result of the very high charge carrier mobility from the surface electrons on the graphene layer and the LSPR absorption from the various TM PNSs, ultrafast charge transfer from the resultant graphene-TM PNSs is facilitated when they get excited by the visible-NIR light source

Received: December 2, 2022
Accepted: January 4, 2023
Published: January 19, 2023

Memories - Materials, Devices, Circuits and Systems 4 (2023) 100052

Contents lists available at ScienceDirect

Memories - Materials, Devices, Circuits and Systems

journal homepage: www.elsevier.com/locate/memori

Design of novel hybrid - digitally controlled oscillator for ADPLL

Mohd Ziauddin Jahangir [a,*], Paidimarry Chandra Sekhar [b]

[a] Department of Electronics and Communication Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad, India
[b] Department of Electronics and Communication Engineering, Osmania University college of Engineering, Hyderabad, India

A R T I C L E I N F O

Keywords:
Hybrid-DCO
Hybrid-VLRO
Hybrid-CCRO
CSRO
High resolution DCO

A B S T R A C T

Digitally Controlled Oscillators (DCOs) are an integral part of All Digital Phase Locked Loops (ADPLLs). It is used to generate output frequency corresponding to the applied digital input. But, due to practical limitations in circuit design, DCOs resolution will be highly limited. Delta Sigma Modulator (DSM) is generally used in DCOs/ADPLLs to obtain a greater resolution. However, using DSM will lead to introduction of frequency spurs in the output spectrum. In this work we propose an alternate method to increase the frequency resolution of DCO without introducing frequency spurs. Novel Hybrid DCO architecture is proposed in this work to increase the resolution. Hybrid DCO proposed in this work has both digital and analog control for tuning frequency. Digital control input of the proposed DCO provides coarse frequency control and the analog control input provides fine frequency control. It has been demonstrated in this work that by integrating Hybrid DCO and DSM with an analog low pass filter (LPF), resolution can be greatly increased, without introducing spurs in the spectrum. As a proof of concept, two Hybrid Digitally controlled Ring Oscillators (DCROs) are designed in 90 nm CMOS process, and their period Jitter performance is compared.

1. Introduction

ADPLLs use DCOs to synthesize signals. A typical ADPLL block diagram is as shown in Fig. 1.

DCOs used in ADPLL can be either an LC based DCO [1] or a Ring oscillator based DCO [2,3]. The earliest implementation of DCO involved integrating DAC with VCOs as can be found in [4]. However, this approach leads to degradation of the noise performance of the resultant DCO as the noise from DAC will be added to already noise VCOs. Improvements in the DCO design was made by direct digital control of the elements in the VCO circuit. [1] uses a digitally switched, capacitor bank to convert an LC- VCO to a DCO. As load capacitance is switched on or off based on digital inputs, it varies the frequency of the oscillator. In [2], the number of inverters, in the ring architectures are switched to vary the frequency of a ring oscillator. In [3] the frequency variations are made possible using a digitally controlled tri-state inverter. Number of inverter driving the node will increase the current and therefore will lead to variable charging times. Also, it can be observed, in [2,5] that a digitally controlled varactor capacitance load is used to convert a ring oscillator into a DCO. Similar, architectures are used in latest DCO like [6,7].

The approaches of designing DCO, listed in preceding paragraph, are not easily scalable. Because increasing range will lead to increase in the area and improving resolution is mostly not possible as it is restricted by the minimum device dimensions.

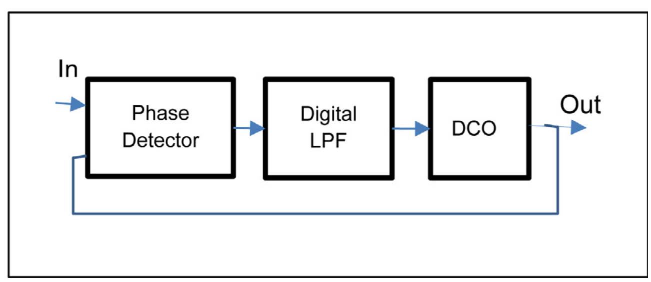


Fig. 1. Block diagram of a typical ADPLL.

The resolution of a conventional DCO usually depends on the number of digital inputs, as shown in Eq. (1).

$$\text{Freq. Resolution} = \frac{F_{\text{Range}}}{2^n} \tag{1}$$

For example, an 8-bit DCO operating with a max frequency range of 2 GHz to 4 GHz would have a resolution of 2G/256 = 7.8 MHz which is inadequate for any real time commercial applications.

Several mechanisms are being used to improve the resolution. A general method is to use a Delta Sigma Modulator (DSM) to increase

* Corresponding author.
E-mail address: jahangir_ece@cbit.ac.in (M.Z. Jahangir).

https://doi.org/10.1016/j.memori.2023.100052
Received 6 December 2022; Received in revised form 3 April 2023; Accepted 22 April 2023

A 1.2GHz Frequency Range, 153.4 dBc/Hz FoM, Low Phase Noise, Current Starved Multi-Path Ring VCO

Mohd Ziauddin Jahangir[1] and Paidimarry Chandra Sekhar[2]

[1]Mohd Ziauddin Jahangir, Department of Electronics and Communication Engineering, CBIT, Hyderabad, India
[2] Paidimarry Chandra Sekhar, Department of Electronics and Communication Engineering, Osmania University College of Engineering, Hyderabad, India
e-mail: md.ziauddin.jahangir@gmail.com

Abstract—**This article describes the design of a low power, low phase noise, multi-Path ring VCO (MPRVCO) in 90nm CMOS process. The proposed oscillator achieves a tuning range of 1.2GHz operating at 1.65 GHz center frequency, with reduced phase noise of -87.3dBc/Hz at 1MHz offset. The proposed MPRVCO achieves a FoM of 153.4dBc/Hz, consuming 0.657mW power at 1.65GHz frequency. The proposed VCO utilizes the Current starving technique for frequency variation. Sub-threshold transistor is used to obtain monotonic and linear frequency tuning characteristics. The proposed VCO is one of the very few Current Starved ring VCOs capable of producing such low phase noise in 90nm CMOS process while operating in GHz frequency range.**

Index Terms— **Multipath VCO, Multi-Path ring VCO, Low Phase Noise Ring Oscillator, Linear frequency characteristic.**

I. Introduction

Ring Oscillators are preferred choice for VCO, for applications requiring smaller area and low power consumption. A conventional ring oscillator is shown in the Fig 1. It consist of three or more, odd number of inverters, connected in a ring. With the recent trend of implementing PLLs digitally, there has been a great focus in development of ring oscillators.

Different structures of ring oscillators have been explored in recent years, to obtain better performance. Ring Oscillators are designed using inverters followed by latches, in [1]. The designed VCO increased the frequency of oscillations and reduced the phase noise . In [2] it was analyzed that there exist multiple path for the oscillator and hence there existed many modes of oscillation and conditions for operating at higher oscillation frequency mode was derived. Further modification was done in [3] to stabilize the frequency. It can also be observed that the concept of injection locking is also being used to improve the frequency of the ring oscillator. Using injection, the phase noise of the oscillator was reduced in [4]. Similarly, we find use of multi path ring oscillator as another technique for implementing VCO as can be seen in [5][6]. It is also being used for implementing TDCs, as can be found in[7].

In summary, multi-Path and injection locking are two techniques that can be used to increase the frequency range and reduce the phase noise of ring oscillators. However, Multi-path technique is quite preferable, as it is simple to implement when compared to injection locking. A simplified block diagram of the multi-path ring oscillator is shown in the Fig 2. There exist various modes of oscillation in multi-path ring VCO [2]. Appropriate mode of oscillation can be selected by appropriately sizing the transistor of given input. Though there are many multi-path ring oscillators already proposed in the literature, like [2][5], the mechanism of its frequency control is usually, by means of controlling its capacitive load. This mechanism yields non-linear frequency characteristics. Very few literature exist on the current based frequency control of MPRVCO and none describing the effect of MOSFET's channel length on the MPRVCO.

The main contribution of this work is, it presents a current-starved multi-Path ring VCO, designed using long channel transistors. The longer channel transistors provide relatively less noise contribution which results in the reduction of the Phase Noise. Further, the use of sub-threshold MOSFET as current source will result in lower power consumption and better linearity.

The article is organized as follows. The motivation behind the development of the proposed VCO is discussed. Later, the circuit of the proposed CS-MPRVCO is described. Then the effect on the phase noise and frequency tuning is described. Lastly, the simulation results are presented.

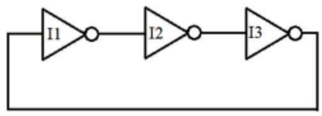


Fig 1 A conventional Ring Oscillator

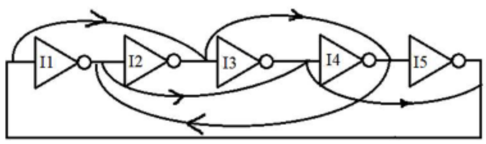


Fig 2 Illustrative diagram of a multi-path ring oscillator

II. Motivation and Design

The phase noise of a CMOS ring oscillator is usually very difficult to reduce as it has very stringent trade-off with parameters like W/L, Power, etc. As per [8], it is not dependent on the number of stages N. Hence, architectural changes are required to reduce the phase noise of ring oscillators. The below discussion gives an idea on conventional solution to phase noise reduction and its limitations.

A. Phase Noise of an oscillator

The Single side band phase noise of an oscillator is as shown in the Fig 3. It consists of three parts. The noise at frequency

Digital Object Identifier 10.29292/v18i1.676

Aerospace Systems (2023) 6:595–604
https://doi.org/10.1007/s42401-023-00222-9

ORIGINAL PAPER

Ionospheric irregularities measurement using Indian SBAS-GAGAN

Kaitha Praveena[1] · Perumalla Naveen Kumar[1] · D. Krishna Reddy[2]

Received: 15 December 2022 / Revised: 13 April 2023 / Accepted: 19 April 2023 / Published online: 2 May 2023

Abstract

In order to improve the performance of a navigation systems, scintillation studies and Ionospheric Total electron content (TEC) are important. The amplitude scintillation index (S_4), S_4 corrections and Rate of change of TEC index (ROTI) parameters are analysed for different seasons. For the analysis Visakhapatnam station (Lat: 17.78, Long: 83.22) and Lucknow (Lat: 26.76, Long: 80.88) station, GAGAN receiver data for the year 2016 is considered based on four quiet days and four disturbed days and consider the highest Kp index values for Visakhapatnam station and Lucknow station. This work shows the variation of S_4 index and ROTI parameter variation during different seasons for both the stations. The correlation coefficient (CC) of S_4 index and ROTI are presented. The results show that the CC are high for disturbed days compared to the quiet days for both the considered stations. For Lucknow station, it is observed that CC values are high compared to the Visakhapatnam station.

Keywords ROTI · S_4 index · GAGAN · TEC

Abbreviations

GPS	Global positioning system
ROTI	Rate of change of TEC index
GAGAN	GPS aided geo augmented navigation
AAI	Airports Authority of India
ISRO	Indian space research organization
CC	Correlation coefficient
S_4	Amplitude scintillation index
TEC	Total electron content
ROT	Rate of change of TEC
SNAS	Satellite navigation augmentation system
SI	Signal intensity
C/N_0	Carrier to noise ratio

1 Introduction

There are many errors, such as Ionospheric scintillations, affect the performance of GPS. A standalone GPS does not work well for certain applications, including aircraft landings, for example. Accordingly, the Airport Authority of India (AAI) and the Indian Space Research Organization (ISRO) developed a Satellite Based Augmentation System (SBAS) known as GPS Aided Geoaugmented Navigation System (GAGAN) [1]. Understanding of the Ionospheric environment is essential for the successful implementation of the GAGAN system in India [2]. Many researchers are worked for understanding the ionospheric irreguritiesin the context of GAGAN. Ionospheric scintillations can cause a GNSS receiver's performance to suffer, ranging from a reduction in location accuracy to a full loss of lock. In order to provide uninterrupted positioning services, it is helpful to identify and mitigate undesirable ionospheric scintillation effects on propagation pathways [3]. Carrier to Noise ratio (C/N0) sampling is required for amplitude scintillations, while carrier-phase observations are needed for phase scintillation [4, 5]. Since the ionospheric irregularities behavior is a random phenomenon, statistical parameters are required to characterize the randomness. The rate of change in TEC, it is represented by ROT and may also be obtained via a GNSS receivers, and also used to measure Ionospheric anomalies. Ionospheric abnormalities have a random nature, and statistical factors are important in understanding the randomness [6]. Using ROTI observations, it is possible to predict the occurrence of scintillations that are responsible for those irregularities [7, 8]. On the other hand, amplitude scintillations may result in a reduction of the

✉ Kaitha Praveena
kaithapraveena@osmania.ac.in

1 Advanced GNSS Research Laboratory, Department of ECE, University College of Engineering, Osmania University, Hyderabad, India

2 Department of ECE, CBIT, Hyderabad, India

Automated Identification of Brain Tumor using Image Transformation Methods

N. Phani Bindu
Research Scholar
Dept. of Electronics and Communication Engineering
University College of Engineering(Autonomous)
Hyderabad-500007, Telangana, India
e-mail: phani.bindu@gmail.com

Dr. P. Narahari Sastry
Professor
Dept. of Electronics and Communication Engineering
ChaitanyaBharathi Institute of Technology,Gandipet
Hyderabad, Telangana, India
e-mail: naraharisastry_ece@cbit.ac.in

Article Info
Page Number: 8061 - 8077
Publication Issue:
Vol 71 No. 4 (2022)

Abstract

In medical image processing, automated detection and classification of Brain Tumor Images (BTI) is very important. Tumors are nothing but the abnormal cells which grow in the brain and directly affect all the healthy cells. In young generation the effect of brain tumor is rapidly increasing. Manual detection and classification of brain tumors can cause human errors. Automated detection and classification of tumor is required as it reduces the burden of human observer and the accuracy also will not be affected due to use of large number of images. Accurate detection and classification of the tumors is required for diagnosis and subsequent treatment planning .Generally, electronic equipment is used in brain tumor diagnosis. The efficient and most popular technique used for diagnosing the brain tumor is Magnetic Resonance Imaging (MRI).This paper uses an image transformation technique named Discrete Cosine Transform (DCT) to obtain the test data results as normal or abnormal images by using the trained dataset images and calculate the percentage of accuracy, sensitivity and specificity using the confusion matrix attributes. Then the obtained abnormal images are further classified by a novel method of segmentation. In this process Otsu's Binirization technique is used to obtain the binary transformation of an image and clustering algorithm named k-means clustering is used to segment the required area of an image. A Discrete Cosine Transform (DCT) is employed for obtaining the features of the image and these extracted features are given to kernel SVM and the Cross validation method is used for enhancement and SVM generalization to classify the Benign and Malignant tumors. These methods are helpful for early detection and also assist doctors in identifying the severity of tumor.

Article History
***Article Received**: 25 March 2022*
***Revised**: 30 April 2022*
***Accepted**: 15 June 2022*
***Publication**: 19 August 2022*

Keywords-Brain Tumor; DCT; Magnetic Resonance Imaging (MRI); Otsu Binirization technique; SVM Classification

Soft Computing (2023) 27:9179–9189
https://doi.org/10.1007/s00500-023-08420-5

FOCUS

Automated brain tumor detection and segmentation using modified UNet and ResNet model

N. Phani Bindu[1] · P. Narahari Sastry[2]

Accepted: 4 May 2023 / Published online: 17 May 2023

Abstract
The early automatic detection of brain tumors in MRI scans is a challenging endeavor due to the high resolution of the images. For a very long time, continual research efforts have been floating a new notion of substituting various grayscale anatomic parts of diagnostic pictures with suitable colors. If successful, this would be an effective way for radiologists to circumvent the challenges they now encounter. The coloring of grayscale photos is a complex process that aims to improve the contrast of different sections of an image by changing grayscale images into color pictures with high levels of contrast. It is common for the predictions to be lacking in fine detail when simply a U-Net design is used; to assist alleviate this issue, cross or skip connections may be introduced between the blocks of the network. Instead of creating a skip connection every two convolutions as it now is in a ResBlock, the skip connections cross from a portion of the same size in the downsampling route to a part in the upsampling path. This improves the overall accuracy of the model and performs better when compared to traditional UNet model.

Keywords Brain tumor segmentation · UNet · ResNet · Encoder · Decoder

1 Introduction

Research in the field of medical imaging has led to the creation of many different types of diagnostic procedures. Each comes with its own set of benefits and drawbacks. In the field of medicine, “medical imaging” refers to the process of producing scans of the human body with the purpose of assisting in the diagnosis of a medical ailment. Not only does it help in the diagnosis and treatment of illness, but it also makes it possible to analyze the diseases present in the patient. This is an extremely beneficial aspect of this technique. It does this by comparing the subject’s anatomy and physiology. In Magnetic Resonance Imaging (MR)I imaging, one of the most important topics to discuss is the segmentation of brain tumors. MRI image segmentation is the process of dividing the image into pixel groups based on the intensity and texture values. The goal of this process is to simplify the process of analyzing the image (Clark et al. 2013).

An MRI scan is used most often in the diagnosis of brain malignancies. The MRI machine performs its functions in the same manner. A radio wave that is produced by the patient’s body is picked up by an antenna (coil) while the patient is being scanned by a radio transmitter. The radio wave is put through a series of computations by a computer program, which ultimately leads to the production of a picture of magnetic resonance (MRI).

Primary tumors and secondary tumors are the two distinct classifications that may be used to cancerous growths (Lin et al. 1991). Malignant tumors spread slowly and persist. They are fast increasing, although the boundaries of their territory are not entirely apparent. There is a possibility of developing primary as well as secondary cancers (Khambhata 2016; Singh and Ahuja 2019; Kaur

✉ N. Phani Bindu
phani.bindu@gmail.com

P. Narahari Sastry
naraharisastry_ece@cbit.ac.in

[1] Department of Electronics and Communication Engineering, University College of Engineering Osmania University, Hyderabad, Telangana 500007, India

[2] Department of Electronics and Communication Engineering, Chaitanya Bharathi Institute of Technology, Osmania University, Osman Sagar Rd, Kokapet, Gandipet, Telangana 500075, India

Materials Today: Proceedings 72 (2023) 465–470

Contents lists available at ScienceDirect

Materials Today: Proceedings

journal homepage: www.elsevier.com/locate/matpr

Design and evaluation of nanostructured formulations of rosuvastatin

N. Sadhana [a,*], B.V.S. Praveen [b,*], G. Jyothi [a], P. Vishnu [a,*]

[a] Department of Pharmaceutical Engineering, B V Raju Institute of Technology, Narsapur, Medak 502 313, India
[b] Department of Chemical Engineering, Chaitanya Bharathi Institute of Technology (CBIT), Hyderabad 500075, India

ARTICLE INFO

Article history:
Available online 6 September 2022

Keywords:
Rosuvastatin
Nanosponges
Release Kinetics
Characterization
In-vitro drug release

ABSTRACT

Cyclodextrins-based nanocarriers are versatile formulations used for drug delivery applications. These nanosponge formulations are hyper-cross-linked polymer structures used to improve the solubility and bioavailability of various poorly soluble drugs. Rosuvastatin is an anti-hyperlipidemic drug that acts by inhibiting of β-hydroxy β-methyl glutaryl-CoA (HMG CoA) reductase enzyme, which is essential for cholesterol synthesis, that is used to treat hyperlipidemia, a condition that has elevated cholesterol levels. In the present study, Rosuvastatin was synthesized by the emulsion solvent evaporation method to prepare nanosponges using polymers ethylcellulose and β-cyclodextrin and evaluated for particle size, and surface morphology using scanning electron microscopy, zeta potential, drug-excipient compatibility using Fourier Transmission Infrared Spectroscopy (FTIR). Apart from these, entrapment efficiency, drug release, and release kinetics were also studied. Nine formulations using different concentrations of polymers were prepared. The best formulation (R showed the average particle size was 386 nm, entrapment efficiency calculated showed 75.45, and drug release was 99.77 % at 8 h drug release kinetics showing zero-order drug release and Higuchi diffusion mechanism.

Selection and peer-review under responsibility of the scientific committee of the International Conference on Novel Materials and Technologies for Energy and Environment Applications (NMTEEA-2022).

Materials Today: Proceedings 72 (2023) 286–292

Contents lists available at ScienceDirect

Materials Today: Proceedings

journal homepage: www.elsevier.com/locate/matpr

A review on nano composite polymer electrolytes for high-performance batteries

N. Laxmiprasanna [a], P. Sandeep Reddy [a], G. Shiva Kumar [a], M. Balakrishna Reddy [a], Kiran Kumar Ganta [b], Venkata Ramana Jeedi [b,*], B.V.S. Praveen [c,*]

[a] Department of Chemical Engineering, B V Raju Institute of Technology, Narsapur, Telangana 502313, India
[b] Department of Physics, B V Raju Institute of Technology, Narsapur, Telangana 502313, India
[c] Department of Chemical Engineering, Chaitanya Bharathi Institute of Technology (CBIT), Hyderabad 500075, India

ARTICLE INFO

Article history:
Available online 2 August 2022

Keywords:
Nanocomposites
Polymer electrolyte
Plasticizer
Nanofiller
Dopants

ABSTRACT

Polymer electrolytes (PEs) are lightweight, flexible, non-flammable solutions and provide a feasible way to replace organic liquid electrolytes because of the safety issues faced by energy storage devices. Inorganic nanofillers are vital in enhancing ionic conductivity by creating additional pathways in polymer-based nanocomposite solid electrolytes. The advances in polymer electrolytes play an essential role in developing the energy storage devices such as batteries. This article briefly summarises the preparation techniques of different polymer electrolytes, developing and focusing on fundamental properties such as surface morphology, electrolyte phases, and the influence of dopants, fillers, and plasticizers on the ionic charge transfer, ionic conductivity, dielectric properties, and transport properties using impedance spectroscopy. We anticipate that gives clear direction for developing energy storage devices.

Selection and peer-review under responsibility of the scientific committee of the International Conference on Novel Materials and Technologies for Energy and Environment Applications (NMTEEA-2022).

Journal of the Indian Chemical Society 100 (2023) 100863

Contents lists available at ScienceDirect

Journal of the Indian Chemical Society

journal homepage: www.journals.elsevier.com/journal-of-the-indian-chemical-society

Kinetic modelling and simulations studies for propanoic acid esterification process

Mallaiah Mekala[a,*], Raju Kalakuntala[b], Srinath Suranani[b]

[a] Department of Chemical Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad, 500075, India
[b] Department of Chemical Engineering, National Institute of Technology, Warangal, 506004, India

ARTICLE INFO

Keywords:
Propanoic acid
Kinetics
Model
Simulation

ABSTRACT

The esterification of propanoic acid with n-butanol to produce n-butyl propionate and water in the influence of an Amberlite catalyst was investigated using a batch reactor. The catalyst was chosen based on its ability to perform in this reaction. The temperature of the reaction runs between 363.15 K and 403.15 K, while the propanoic acid to n-butanol molar ratio is around 1:1 and 1:4. The selection of the catalyst loading is dependent on the volume of the reaction mixture. The catalyst dosage was kept within a 1%–3% by weight range. The kinetics of conversion have been researched in relation to the reaction temperature, mole ratio, catalyst size, stirrer speed, and catalyst quantity. The catalyst dosage and reaction temperature, according to the study, have a substantial influence on how soon the system achieves equilibrium. The pseudo homogeneous kinetic model is developed and tested against experimental results. Under the specified conditions, model predictions and empirical observations accord well. Arrhenius equation was used for calculation of rate constants and energy of activation. The forward reaction's frequency factor & activation energy are 18.554 L/mol.min and 30350 J/mol, correspondingly. Since the equilibrium constant increases as the temperature rises, the reaction is endothermic.

LANGMUIR

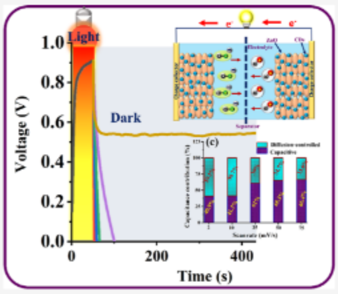
pubs.acs.org/Langmuir Article

N-Doped Carbon Dots and ZnO Conglomerated Electrodes for Optically Responsive Supercapacitor Applications

Rupam Sinha,* Nirmal Roy, and Tapas K. Mandal

Cite This: https://doi.org/10.1021/acs.langmuir.3c00300

ABSTRACT: The over-dependence of human society on fossil fuels for energy is exhausting the level of such non-renewable energy sources. Alternative energy storage systems have gained more popularity recently to counter this issue. In this context, we report the fabrication of N-doped carbon dot (N-CD)-decorated ZnO-based electrodes for supercapacitor applications. Due to the light-responsive nature of the N-CDs and ZnO, the electrode was also responsive under the influence of UV light. After the experimental tests, it was found that the areal capacitance value of the supercapacitor increased upto ~58.9% when illuminated compared to that under the dark conditions. Moreover, the device showed a maximum areal capacitance of 2.6 mF/cm^2 after photocharging and galvanostatically discharging at a current density value of 1.6 μA/cm^2, which is quite comparable with the previously reported data. The doping of N-CDs with ZnO showed a significant improvement in the areal capacitance value under both illuminated (~58.64%) and dark conditions (~22.08%) compared to the case of pristine ZnO, which justifies the purpose of attaching N-CDs with ZnO. Therefore, in brief, we have fabricated a photoresponsive electrode material for supercapacitor application by combining N-CDs and ZnO. An explicit electrochemical characterization of the electrode was also done to identify the contribution from diffusion-controlled capacitance and double layer capacitance, and it was observed that the diffusion-controlled capacitance gets reduced from 59.1 to 33.6% when the scan rate is increased from 2 to 75 mV/s. Moreover, a detailed study has also been done to understand the reaction mechanism. It was confirmed that the defects in the electrode material played a vital role in the intercalation of K$^+$ ions.



materialstoday: PROCEEDINGS

Volume 72, Part 1, 2023, Pages 336-339

Simulation of pyrolytic conversion of Walnut shell waste to value added products

Bhushan Goklani [a], P.V. Naga Prapurna [a], S. Srinath [b]

https://doi.org/10.1016/j.matpr.2022.08.024

Abstract

materialstoday: PROCEEDINGS

Volume 92, Part 2, 2023, Pages 458-464

Sustainable biodiesel: A comprehensive review on feedstock, production methods, applications, challenges and opportunities

Madhuri Pydimalla, Sadia Husaini, Akshara Kadire, Raj Kumar Verma

https://doi.org/10.1016/j.matpr.2023.03.593

Reaction Chemistry & Engineering

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal

Check for updates

Cite this: DOI: 10.1039/d2re00543c

Production of biodiesel: kinetics and reusability studies of the Mg–Al hydrotalcite catalyst using Jatropha oil

B. V. S. Praveen,[a] Narayan C. Pradhan,[b] Anup Ashok,[c] Ramesh Kumar Guduru,[d] Rakesh Kumar Vij[e] and Lakshmana Rao Jeeru[*e]

Received 12th December 2022,
Accepted 3rd April 2023

DOI: 10.1039/d2re00543c

rsc.li/reaction-engineering

This study focuses on the use of the co-precipitation method to prepare Mg–Al layered double hydroxide (LDH) catalysts for the transesterification conversion of Jatropha oil into biodiesel. X-Ray diffraction (XRD), scanning electron microscopy (SEM), and BET methods were used to characterize the crystallinity, morphology, and surface area of the catalyst. The effects of varying reaction conditions, such as time, catalyst loading, reaction temperature, and methanol to oil molar ratio, on biodiesel yield, were investigated. The optimal experimental conditions for achieving a 95 percent yield were identified as a temperature of approximately 70 °C, a catalyst loading of 3 percent by weight, and a molar ratio of methanol to Jatropha oil of 12 to 1. Catalyst regeneration was also studied to determine reusability, and the activation energy was evaluated to be around 26.7 kJ mol^{-1}. The results showed that the catalyst could be reused for up to 5 cycles with a yield reduction of less than ~10%. The study demonstrates the potential of using inexpensive Mg–Al LDH catalysts for large-scale biodiesel production.

Introduction

As the global population continues to rise and the demand for energy increases, researchers have been compelled to explore various alternatives to fossil fuels. While energy can particulate matter, and fewer unburned hydrocarbons. Its high flash point can be used in its pure form or blended with other fossil fuels (such as B20), making it safer to transport and handle than petroleum.[6–8]

Researchers have developed several alternatives to fossil

e-ISSN 2320 –7876 www.ijfans.org
Vol.12, Iss.1, Jan 2023
Research Paper , UGC CARE Listed (Group -I) Journal

SIMULATION STUDIES FOR PRODUCTION OF ISOPROPYL ACETATE USING ASPEN PLUS

Sk.Shaheda[1], Prof.V.V.BasavaRao[2],Dr. M. Mukunda Vani[3]

[1]Assistant Professor, Department of Chemical Engineering, Anurag University Venkatapur, Ghatkesar, Telangana

[2]Professor, Department of Chemical Engineering, University Technology, Osmania University, Telangana

[3]Professor,Department of Chemical Engineering ,Chaitanya Bharathi Institute of Technology, Telangana

Address for correspondence : shaheda.agi@gmail.com

Abstract: Isopropyl acetate is one of the most important fine chemical intermediate, which is formerly given as universal solvent. Acetic acid with Isopropyl alcohol to produce Isopropyl in amberlyst 36 wet in different reactors using Aspen plus has been researched. The feed mole ratio was differs from 1:1 to 1:1.6 and reaction temperatures were differs from 333.15 to 358.15 K for the solid catalysts. The acetic acid conversion was established to grow more in continuous stirred tank reactor compared to Gibbs reactor for a conversion is 61% under 85 ^{0}C for the feed of 1: 1.2 of Isopropyl Alcohol and acetic acid in the chosen method. In present work further studies was done for the combination of the continuous stirred tank reactor connected to simple purification pillar for the outcome of Isopropyl acetate was simulated using Aspen Plus to get high purity has been considered. The results were achieved with high conversion and purity of product 90% at feed temperature 358.15k with the reflux ratio was 5 and number of stages 22.

Key words: Amberlyst 36, Gibbs reactor, continuous stirred tank reactor, distillation column, chosidel method.

I.INTRODUCTION

Acetic acid (HoAC) esters are a usual application1 of reactive purification, effective to avoid the response equanimity conditions by extract many of the outcomes from the reaction location. You and colleagues2 considered uninterrupted RD

subsequently, at 1 at m) acquired in the liquid stage.

The mentioned 3 methods have a treble minimal -boiling heterogeneous azeotrope consists of different quantity of the alcohol

Chemical Engineering Research and Design 196 (2023) 451–467

Available online at www.sciencedirect.com

Chemical Engineering Research and Design

journal homepage: www.elsevier.com/locate/cherd

Dynamic modeling and optimal control of reactive batch distillation: An experimental case study

B. Ganesh[1], K.Yamuna Rani*

Process Dynamics & Control and Artificial Intelligence Group, Process Engineering & Technology Transfer Department, CSIR, Indian Institute of Chemical Technology, Hyderabad 500007, India

ARTICLE INFO

Article history:
Received 14 March 2023
Received in revised form 24 June 2023
Accepted 28 June 2023
Available online 30 June 2023

Keywords:
Reactive batch distillation
Experimental study
Dynamic modeling
Start-up period
Genetic algorithm
Optimal control
Esterification reaction
Set point trajectories tracking

ABSTRACT

In the present study, dynamic modeling, optimal operation and control of a reactive batch distillation process is illustrated based on experimental and simulation studies using methyl acetate production case study. An equilibrium stage model, incorporating non-ideal VLE and start-up region of heating, is developed based on data generated on an experimental unit by identifying five input parameters representing uncertainties. The optimal control problem is solved using genetic algorithm, considering the objective function identified on the basis of trend analysis using the developed dynamic model, to find the optimal reflux ratio, heat input to the reboiler, and mole ratio of methanol to acetic acid in the initial reaction mixture. Open-loop and closed-loop implementations clearly illustrate the improved performance with respect to the quantity of methyl acetate in the distillate product with a reasonably high conversion and product purity within reasonably short batch duration, illustrating the successful optimal control implementation experimentally.

materialstoday: PROCEEDINGS

Volume 76, Part 1, 2023, Pages 166-169

Fuel oil from plastic waste

Prasad Babu Koorla, Narayana Dasari, Noorullah Shareef Mohammad

Chemical Engineering Department, Chaitanya Bharathi Institute of Technology, Gandipet, Hyderabad 500075, India

Available online 16 November 2022, Version of Record 16 February 2023.

What do these dates mean?

Chapter 9

Print ISBN: 978-81-959996-2-0, eBook ISBN: 978-81-959996-6-8

Experimental Study and Performance Enhancement of Surface Coating Techniques

K. Samhitha Reddy [a++], V. Shreya Reddy [a++], J. Sujitha [a++] and Madhuri Pydimalla [a#*]

DOI: 10.9734/bpi/taier/v6/4840A

ABSTRACT

Surface coatings have changed over time to accommodate the changing demands of the processing sector. Powder coatings are desirable due to their adaptability, high performance and minimal environmental effect. This study aims to investigate the role of powder coatings in the evolution of conventional surface coating technology for practical applications in the field of metal coating, particularly with regard to office and home appliances. The primary objective of this paper is to provide a thorough analysis of the benefits and drawbacks of polymer powder coatings versus liquid-based coatings. The second objective was to compare polymer coatings to additive-filled powder coating. In this study, conventional substrates including copper, aluminum, galvanized iron, brass, cement plank, and wood block were utilized. For the priming coat, the materials were first dry scuffed and then immersed in a 3 in 1 chemical (a combination of zinc phosphate and magnesium phosphate). Coating the surfaces of the prepared substrates with spray-gunned liquid paint (on one side of the panel) and

ScienceDirect

Sustainable Energy Technologies and Assessments

Volume 56, March 2023, 103079

An enhanced opportunistic rank-based parent node selection for sustainable & smart IoT networks

Premkumar Chithaluru [a g], Aman Singh [b h], Mahmoud Shuker Mahmoud [c], Sunil Kumar [d g], Juan Luis Vidal Mazón [b g i], Ahmed Alkhayyat [e], Divya Anand [f b]

Show more

Share Cite

https://doi.org/10.1016/j.seta.2023.103079

Get rights and content

Abstract

The Internet of Things (IoT) is a network of interconnected devices that includes low-end devices (sensors) and high-end devices (servers). The routing protocol used the Low-Power and Lossy Networks (RPL) protocol, which was designed to collect data in Low-Power and Lossy Networks (LLN) efficiently and reliably. The RPL rank property specifies how sensor nodes are placed in Destination Oriented Directed Acyclic Graphs (DODAG) based on an Objective Function (OF). The OF includes information such as the Expected Transmission Count (ETX) and packet delivery rate. The rank property aids in routing path optimization, reducing control overhead, and maintaining a loop-free topology by using rank-based data path validation. However, it causes many issues, such as optimal parent selection, next-hop node selection, and network instability. We proposed an Enhanced Opportunistic Rank-based Parent Node Selection for Sustainable and Smart IoT Networks to address these issues. The optimal parent node is determined by forecasting the expected energy of each node using Received Signal Strength (RSS) and an enhanced reinforcement learning algorithm. The proposed method addresses the issue of selecting the next-hop neighbor node and improves routing stability. Furthermore, when a large number of new nodes try to join the sustainable IoT-based smart cities, the proposed technique provides optimal load balance.

Introduction

An IoT is a collection of network devices that are used to communicate and collect information from monitoring areas throughout the wireless node's links[1]. The data packets can be transmitted via multiple sensor nodes with a single gateway node, and it is linked to other networks such as wireless Ethernet (WLAN)[2]. It contains many sensor nodes and base stations to communicate with the monitored area. The environmental or Physical conditions are monitored by using the IoT[3], the conditions may be in the form of sound, temperature, pressure, wind speed, and direction, pollution levels, humidity, etc and this monitored information will be passed via the network to a central location[4].

The IoT network has been constructed first then its applications can be used in several domains such as military, sports, transportation, healthcare, and other industrial fields[5]. In general, an IoT is composed of a significant number of sensor nodes ranging from hundreds to thousands[6]. In IoT, some sensor node types of equipment are exploited, like a microcontroller and radio transceiver, along with an antenna, an energy source, an interfacing electronic circuit, and a battery[7]. The sensor node's size will range from the shoe box's size to as small as the grain of dust's size[8]. The prices of IoT devices have also ranged from a few pennies to hundreds of dollars, depending on the functionality parameters of the sensors, such as computational speed rate, energy consumption, memory, and bandwidth[9].

RPL is an LLN that can be utilized as a routing protocol in IoT with low power consumption[10], but it is vulnerable to packet loss during packet transmission[11]. Depending on the distance vectors, it will act as a proactive protocol and can work on IEEE 802.15.4[12]. It is the best solution for many-to-one communication and assists with one-to-one messages[13]. Routing must frequently change in most cases, while the sensor node only covers a limited range. Every wireless link has been influenced by its environment. Wherever it operates, such fixtures provide combination and reflection, and links are generated either permanently or temporarily undependable[14].

RPL protocol is connected to a Directed Acyclic Graph (DAG)[15], which determines the topology. In a wireless network, each sensor has a rank that increases as the nodes move away from the sink, called Destination Oriented Directed Acyclic Graph (DODAG)[16]. In the said case, the route selection requirement process follows, and by using the lowest range values, nodes can re-transmit the packets[17].

The RSS can be estimated by using the transmission power, the radio capabilities, and the distance between receiver & transmitter[18]. The RSS will be estimated for each received packet. The Received Signal Strength Indicator (RSSI)[19] is calculated using estimated RSS energy. To some extent, the RSSI can be obtained by the Network (NWK) layer, Media Access Control (MAC)[20], and Application Layer (APL)[21] layers. The Link Quality Indicator (LQI)[22] will be made the most convenient way to use the RSSI as a link quality indicator. In addition, the RSSI has developed a method for determining location. The RSSI location-estimation system for IoT entity nodes will be implemented without the use of additional hardware[23].

The issues of the RPL protocol are as follows:

- For LLNs, the RPL has been constructed, and the routing has been executed in a disseminated manner. Conversely, the decisive load balance characteristic has been missed in the RPL protocol[24]. Without load balancing, LLNs can cause a load imbalance for sensor nodes that have many more neighboring nodes than others with uneven data traffic.
- The energy reduction of sensor nodes can be much faster than their light workload. This can cause holes and gaps in the network, causing it to become disconnected[25].
- Furthermore, the RPL protocol only recommends the ETX and Packet Delivery Ratio (PDR) as metrics, and clustering communication reliability is only focused between two nodes[26].
- Under heavy traffic conditions, a perfect connection between two nodes cannot function well because wireless communication can be interfered with merely by the broadcasting of adjacent nodes[27].

The IoT is highly dynamic and sensitive to energy consumption, and the clustering problem becomes a complex task. Energy saving is an important factor when designing the operation of an IoT to gain an extended lifetime. Thus, awareness of the location of the deployed nodes, residual energy, and power consumption of various components of the node is extremely significant. It is critical to research and select the best algorithm for organizing the operation of the IoT in terms of energy savings. These factors are taken into consideration and are investigated in this article.

The main goals of this work are to perform effective parent node selection and extend the IoT lifetime using a novel technique to manage energy consumption. The optimal load balance techniques utilized in this work efficiently solve the problem. In this paper, an enhanced opportunistic rank-based parent node selection clustering protocol is designed, analyzed, and tested for sustainable IoT networks.

The major contributions of the paper are as follows:

- For optimal parent node selection, an Improved Loop-less uni-Directional towards the Root node (ILDR) construction is proposed.
- The proposed method addresses the issue of child node count while also achieving partial routing stability. Furthermore, when an uncontrolled number of new nodes attempt to join the network, the proposed technique provides optimal load balance.
- The performance of the proposed method compared with peer existing methods such as Multilayer Threshold Cluster-Based Energy-Efficient Low-Power and Lossy Networks for Industrial Internet of Things (MTCEE-LLN)[3] and Adaptive Ranking Fuzzy-based Energy-Efficient Opportunistic Routing (ARFOR)[5] protocol in terms of energy consumption, reliability of nodes, beacon transmission, and packet loss.

The structure of the paper is stated as follows: Section “Literature Review” describes some traditional supervised and unsupervised learning-based routing protocols. The methodology of the ILDR protocol is discussed in Section “Proposed methodology”. The efficiency of ILDR with peer-routing protocols is compared in Section “Results and Discussion”. Finally, the conclusions and possible future research are proposed in Section “Conclusions & Future Scope”.

Section snippets

Literature review

IoT is becoming more popular as a result of advancements in wireless and embedded systems. Some applications of the IoT include event recognition, monitoring, tracking, disaster administration, examination, and defensive preservation. IoT routing is one of the most researched concerns[28], [29], [30], [31], [32]. However, it can be difficult to work to expand a routing protocol that will be an effective process due to the inherent characteristics of IoTs, such as application precision,…

Proposed methodology

The IoT is a network of dedicated and dispersed nodes used to monitor the remote environment. LLNs are used to represent sensor networks. IoT routing for identifying the best path between nodes and the sink. Routing protocols must meet the LLN requirements. One of the most widely used routing protocols is RPL. The network between the sink and sensor nodes in RPL is built using acyclic tree formation.

To achieve reliability and energy efficiency in an IoT, the proposed model improves the parent…

Results and discussion

The proposed model is analyzed and compared under various performance metrics such as accuracy, energy consumption, reliability of nodes, beacon transmission, and packet loss. There are different simulators such as OMNET++, MATLAB, Qualnet, NS2, NS3, etc. but Contiki is selected for experimentation in this work. Contiki's open-source nature allows for programmable interventions at various stages of usage via the Cooja simulator and Unit Disk Graph (UDG). The proposed work is simulated using the ...

Conclusions & future scope

An energy-efficient, reliable parent choice is needed in RPL. Therefore, this work has provided an energy-efficient parent node selection for real-time monitoring sensing systems in sustainable IoT networks. The enhanced reinforcement learning method is applied to predict the energy-efficient parent nodes in the sustainable IoT. Based on the predicted energy of the parent, the optimum parent was selected using the enhanced RSS method. This method improved the load balance of the RPL-controlled...

CRediT authorship contribution statement

Premkumar Chithaluru: Conception and design of the study, Analysis and/or interpretation of data. **Aman Singh:** Acquisition of data, Revising the manuscript critically for important intellectual content. **Mahmoud Shuker Mahmoud:** Acquisition of data, Analysis and/or interpretation of data, Drafting the manuscript. **Sunil Kumar:** Conception and design of the study, Analysis and/or interpretation of data, Drafting the manuscript. **Juan Luis Vidal Mazón:** Analysis and/or interpretation of data, Drafting...

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper....

Special issue articles   Recommended articles

References (45)

ChengY. *et al.*
An improved key distribution mechanism for large-scale hierarchical wireless sensor networks
Ad Hoc Netw (2007)

ChenS. *et al.*
Indoor temperature monitoring using wireless sensor networks: a SMAC application in smart cities
Sustainable Cities Soc (2020)

LamaaziH. *et al.*
OF-EC: A novel energy consumption aware objective function for RPL based on fuzzy logic
J Netw Comput Appl (2018)

KimH.S. *et al.*
DT-RPL: Diverse bidirectional traffic delivery through RPL routing protocol in low power and lossy networks
Comput Netw (2017)

ChengL. *et al.*
A survey of localization in wireless sensor network
Int J Distrib Sens Netw (2012)

ShitR.C. *et al.*
Location of things (LoT): A review and taxonomy of sensors localization in IoT infrastructure
IEEE Commun Surv Tutor (2018)

ChithaluruP. *et al.*
MTCEE-LLN: Multilayer threshold cluster-based energy-efficient low-power and lossy networks for industrial internet of things
IEEE Internet Things J (2021)

LashkariB. *et al.*
Crowdsourcing and sensing for indoor localization in IoT: A review
IEEE Sens J (2018)

ChithaluruP. *et al.*
An energy-efficient routing scheduling based on fuzzy ranking scheme for internet of things
IEEE Internet Things J (2021)

BeuchatP.N. *et al.*
Enabling optimization-based localization for IoT devices
IEEE Internet Things J (2019)

View more references

Cited by (4)

An Optimized Bio-inspired Localization Routing Technique for Sustainable IIoT Networks & Green Cities
2023, Sustainable Cities and Society

Show abstract

Using the IoT Sustainability Assessment Test to Assess Urban Sustainability ↗
2024, BIO Web of Conferences

An opportunistic energy-efficient dynamic self-configuration clustering algorithm in WSN-based IoT networks ↗
2024, International Journal of Communication Systems

A cooperation federated forwarder selection technique for maximizing network lifetime in wireless sensor network ↗
2023, Sadhana - Academy Proceedings in Engineering Sciences

View full text

remote sensing

MDPI

Article

Underground Water Level Prediction in Remote Sensing Images Using Improved Hydro Index Value with Ensemble Classifier

Andrzej Stateczny [1,*], Sujatha Canavoy Narahari [2], Padmavathi Vurubindi [3], Nirmala S. Guptha [4] and Kalyanapu Srinivas [5]

1 Department of Geodesy, Gdansk University of Technology, 80232 Gdansk, Poland
2 Department of Electronics and Communication Engineering, Sreenidhi Institute of Science and Technology, Hyderabad 501301, India; cnsujatha@sreenidhi.edu.in
3 Department of Computer Science and Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad 500075, India; padmavathiv_cse@cbit.ac.in
4 Department of CSE—Artificial Intelligence, Sri Venkateshwara College of Engineering, Bengaluru 562157, India; nirmala.s.guptha_cse@svcengg.edu.in
5 Department of Computer Science and Engineering, Vaagdevi Engineering College, Warangal 506005, India; srinivas_k@vecw.edu.in
* Correspondence: andrzej.stateczny@pg.edu.pl

Abstract: The economic sustainability of aquifers across the world relies on accurate and rapid estimates of groundwater storage changes, but this becomes difficult due to the absence of in-situ groundwater surveys in most areas. By closing the water balance, hydrologic remote sensing measures offer a possible method for quantifying changes in groundwater storage. However, it is uncertain to what extent remote sensing data can provide an accurate assessment of these changes. Therefore, a new framework is implemented in this work for predicting the underground water level using remote sensing images. Generally, the water level is defined into five levels: Critical, Overexploited, Safe, Saline, and Semi-critical, based on water quantity. In this manuscript, the remote sensing images were acquired from remote sensing images. At first, Wiener filtering was employed for preprocessing. Secondly, the Vegetation Indexes (VI) (Normalized Difference Vegetation Index (NDVI), Normalized Difference Snow Index (NDSI), Infrared index (IRI), Radar Vegetation Index (RVI)), and statistical features (entropy, Root Mean Square (RMS), Skewness, and Kurtosis) were extracted from the preprocessed remote sensing images. Then, the extracted features were combined as a novel hydro index, which was fed to the Ensemble Classifier (EC): Neural Networks (NN), Support Vector Machine (SVM), and improved Deep Convolutional Neural Network (DCNN) models for underground water level prediction in the remote sensing images. The obtained results prove the efficacy of the proposed framework by using different performance measures. The results shows that the False Positive Rate (FPR) of the proposed EC model is 0.0083, which is better than that of existing methods. On the other hand, the proposed EC model has a high accuracy of 0.90, which is superior to the existing traditional models: Long Short-Term Memory (LSTM) network, Naïve Bayes (NB), Random Forest (RF), Recurrent Neural Network (RNN), and Bidirectional Gated Recurrent Unit (Bi-GRU).

Keywords: improved Deep Convolutional Neural Network; improved hydro index; remote sensing; Support Vector Machine; underground water level prediction; Wiener filter

Citation: Stateczny, A.; Narahari, S.C.; Vurubindi, P.; Guptha, N.S.; Srinivas, K. Underground Water Level Prediction in Remote Sensing Images Using Improved Hydro Index Value with Ensemble Classifier. *Remote Sens.* **2023**, *15*, 2015. https://doi.org/10.3390/rs15082015

Academic Editor: Silvia Liberata Ullo

Received: 18 March 2023
Revised: 7 April 2023
Accepted: 7 April 2023
Published: 11 April 2023

1. Introduction

The groundwater is the world's greatest supply of fresh water, and it is crucial for human consumption and general development of an area [1–3]. The water food energy nexus is mostly dependent on groundwater, which accounts for close to 30% of the world's freshwater reserves [4–6]. Globally, groundwater supplies provide nearly half of the world's drinking water. In addition, groundwater is widely utilized in agricultural processes, and demand is rising as a result of population growth, dietary changes, and climate change [7–9]. In order to measure aquifer degradation and to provide key data for developing a groundwater model to manage the resource, monitoring of groundwater withdrawal (pumping) is required [10–12]. Groundwater can reliably supply the necessary amount of high-quality water; therefore, effective water conservation plans are crucial for the long-term use of groundwater. Excessive groundwater extraction and improper aquifer recharging are two major factors in the depletion of groundwater in many regions. To promote effective use and organized management of groundwater resources, precise estimation and prediction of groundwater recharge should be carried out. This makes groundwater potential mapping with the use of yield data crucial. Data on yields include extraction volume and groundwater velocity at several measurement locations. Groundwater yield is influenced by local geological, topographic, and human-made factors, which are connected to groundwater potential [13,14]. The conventional methods in use to calculate groundwater pumping rely on predictions of agricultural water demand, which are generally made using the evapotranspiration and soil model, and surface water availability [15]. Remote sensing and Geographic Information Systems (GIS) present novel chances for hydrogeological research [16,17]. Maps that show the presence of groundwater are deployed to determine the geology, soil, geomorphology flow depth, rainfall, and land use [18]. Understanding that groundwater (shallow) flow is often influenced by precise surface force and characterized by geologic characteristics that may be deduced from surface data is essential to remotely sensing the water in the ground [19]. Topographic driving factors are a need for the conceived groundwater flow system at various sizes. Since remote sensing offers additional opportunities for groundwater detection, it will be much more successful in estimating water flows in the local ground than the topographically driven flow model. The contributions of this paper are specified below:

- Developed a Wiener filter for inverting the blurring and eliminating the additive noise from the acquired remote sensing images. In addition, the Wiener filter is optimal by minimizing the overall mean square error in the process of noise smoothing and inverse filtering;
- Integrated VI, NDVI, NDSI, IRI, and RVI, and statistical features for feature extraction. The extracted features are discriminative in that they decrease the sematic space between the feature subsets that help in improving the performance of underground water level prediction using remote sensing images;
- Proposed an EC model that includes improved NN, SVM, and DCNN for effective underground water level prediction.

The structure of this research is as follows; Section 2 explains existing works on the topic of underground water level prediction; Section 3 explains the concept of feature extraction and ensemble classification. The results and discussions are described in Section 4, and finally, the conclusion is stated in Section 5.

2. Literature Review

Majumdar et al. [20] developed a new strategy for groundwater level prediction that incorporates elements of the water balance with a Machine Learning (ML) algorithm. Here, the sensor data were used, which assessed various aspects of water balancing and land usage at various geographical resolutions, sequential resolutions, and multi-temporal satellite data. Evapotranspiration, moisture, and land cover were some of the remote sensing products. A cutting-edge ML approach, RF, was utilized to overcome these restrictions since combining these sets of data and then connecting them to groundwater flows utilizing physical models were inherently difficult processes. On both the training and testing datasets, the developed model withdrawals had a high level of accuracy. Veluguri et al. [21] planned to add unique hydro indicators that had not yet gained popularity in earlier methodologies. The developed work estimates statistical properties. In addition, VI incorporated "IRI transformation, Kath Thomas Tasseled cap transformation, normalized difference vegetation index, and simple ratio and the statistical and vegetation index were combined to create a unique hydro index". Then, the ensemble method was used for an identification procedure that includes Deep Belief Network (DBN), NN, RF, and SVM. DBN produces the anticipated outcome by concerning certain metrics; the performance of the developed model was effective in comparison to prior models.

Zhengyang et al. [22] investigated the possibility of groundwater remote sensing in the framework of operational and future satellite-based sensors. The best use of remotely sensed data was integrated with ground-based data for GIS and numerical modeling. Further, to properly and effectively support product design and implementation in the mining region, Suganthi et al. [23] conducted research on feature information extraction. The analysis of Hyperspectral Remote Sensing Image (HRSI) technology's detecting capabilities came first. HRSI has several bands and a high spectral resolution. As appropriate, certain bands can be removed to emphasize desired qualities. The spatial and spectral information was fully exploited in accordance with the features of HRSIs and the CNN depending upon Deep Learning (DL) was utilized for extracting the features. On training the features, CNN enables the computer to automatically extract data features.

Hai et al. set out to clearly explain the extant ML schemes used for modeling Groundwater Level (GWL) and milestones attained in this field [24]. Additionally, suggestions for potential upcoming research topics to increase the precision of GWL forecasts and broaden the field's understanding were provided. For mapping potential areas of underground water in the western desert of Iraq, Mezheret al. [25] utilized statistical methods including Evidential Belief Function (EBF) and Logistic Regression (LR). The geographical interaction between groundwater wells and other conditioning elements was used to evaluate the potential of the groundwater regions. The thematic maps produced in this study demonstrated the efficacy of EBF and LR approaches for mapping groundwater potential.

To forecast GWL underneath ecosystems in the United States of America, Melissa et al. [26] employed satellite-based remote sensing. Groundwater decreases more commonly inside Groundwater Dependent Ecosystems (GDEs) located in regions of the state wherein groundwater resource management is lacking. Groundwater losses are most frequent within the GDEs located in those regions where sustained groundwater control is lacking. Shaimaa et al. [27] has offered two pumping scenarios. In the first scenario, it was assumed that over the next 50 years, the existing extraction rates would not change. The effectiveness of the aquifer was reciprocally influenced by geological and human factors. Nevertheless, despite its undeniable significance, groundwater levels in the utilization of water resources are frequently variable and rely on recharge from precipitation infiltration [28].

3. Methods

The steps involved in the proposed underground water level prediction framework are as follows: initially, the Wiener filter was utilized for image preprocessing. Then, the VI indexes (NDVI, NDSI, IRI, and RVI) and statistical features (entropy, RMS, Skewness, and Kurtosis) were used for feature extraction. Finally, an EC model (NN, SVM, and improved DCNN) was used for underground water level prediction. The illustrative depiction of the developed underground water level prediction model is represented in Figure 1.

3.1. Preprocessing

The input image $I(n,m)$ is subjected to a process under Wiener filtering, which aims to calculate a statistical estimate of an unknown image by employing a similar image as an input and filtering that known image to create the estimate as an output. In the Weiner filter, both the blurring and the gaussian type additive noises are instantly reversed. In the process of inverse filtering and noise smoothing, it minimizes the total mean square error so that the Wiener filter can offer the best mean square error performances. The attained preprocessed image is then passed to the next stage for extracting features. It is mathematically defined as the $I(n,m)$ which takes the Discrete Fourier Transform (DFT) to obtain $A(u,v)$. The original spectrum is estimated by taking the product of $A(u,v)$ and

$G(u, v)$. The Wiener filter $G(u, v)$ is in Equation (1), where $H(u, v)$ is the Fourier transform of the point spread function, and $P_s(u, v)$ is the power spectrum of the signal process, and $P_n(u, v)$ of the noise process.

$$G(u, v) = \frac{H^*(u, v) P_s(u, v)}{|H(u, v)|^2 P_s(u, v) + P_n(u, v)} \tag{1}$$

Here, DFT is applied in a long-time series of satellite images (data cube) along with a different time duration (1 year).

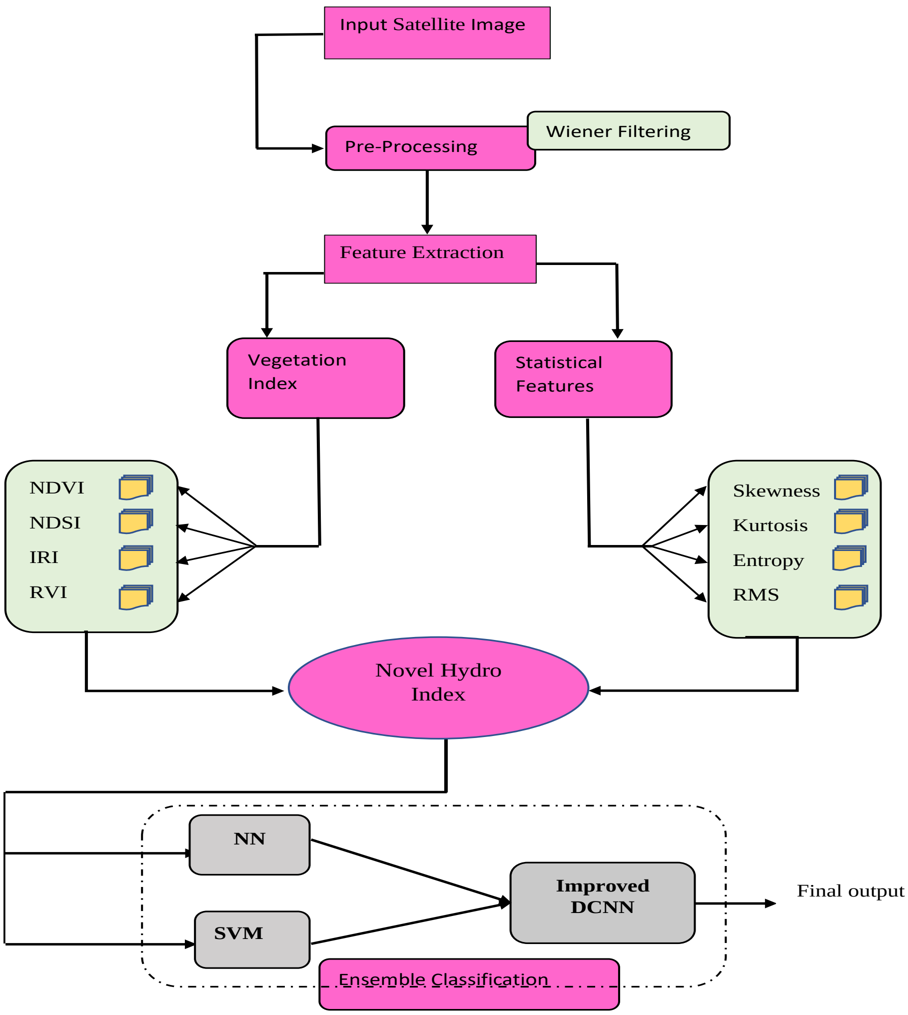


Figure 1. Diagrammatic depiction of the developed underground water level prediction model.

3.2. Extraction of Vegetation Index and Statistical Features

From the preprocessed image, VI indexes and statistical features were extracted. The detailed explanation about the individual VI indexes and statistical features is given below:

NDVI [29]: It is portrayed as a graphical marker that is deployed for analyzing the measurement of remote sensing in a space paradigm, which scrutinizes if the target is noticed or not. The $NDVI$ (ft^{NDVI}) is evaluated as in Equation (2).

$$ft^{NDVI} = \frac{(NIR - R)}{(NIR + R)} \tag{2}$$

Equation (2), NIR and R points out spectral reflectance and Near Infrared (NIR) areas. The $NDVI$ relies on amid values of -1 and +1.

IRI transform: It is portrayed as the proportion of a mixture of *NIR* and Shortwave Infrared (*SWIR*). It is modeled as in Equation (3).

$$ft^{IRI} = \frac{(NIR - SWIR)}{(NIR + SWIR)} \tag{3}$$

NDSI [30]: The NDSI is a measurement of how much the visible (greenish) and SWIR reflectance differences differ from one another. It regulates the variation of 2 bands (one in the NIR or SWIR and the other one in the visible regions of the spectra). This helps to map snow. While most cloud reflectance stays high in the same portions of the electromagnetic spectrum, snow is not only highly reflective in the visible regions of the spectrum, but it is also highly absorbent in the NIR or SWIR regions, allowing for effective removal of most snow and clouds. In Equation (4), band 2 implies visible green with 0.53 to 0.61 µm, and band 5 implies SWIR with 1.55 to 1.75 µm.

$$ft^{NDSI} = \frac{TM_{band2} - TM_{band5}}{TM_{band2} + TM_{band5}} \tag{4}$$

RVI Index: Zyland Kim had [31] suggested a type of RVI [32] as shown in Equation (5).

$$ft^{RVI} = \frac{8\sigma_{hv}}{\sigma_{hh} + 2\sigma_{hv} + \sigma_{vv}} \tag{5}$$

The backscattering coefficients for co-polarization and cross-polarization for various bands are σ_{hh}, σ_{hv}, and σ_{vv}. The RVI scales typically range from 0 to 1 and measure the scattering's unpredictability. The RVI is almost 0 on a smooth, bare surface and rises as the plant develops [33]. The shape, size, and direction of the canopy elements as well as their dielectric qualities all have a role in the intensity of incoming energy that is dispersed by vegetation in the microwave area of the electromagnetic spectrum. Due to their superior spectral and spatial resolution, high-resolution satellite images are frequently employed in groundwater investigations. Additionally, the sensor setup, including the frequency, polarization, and incidence angle, has an impact on the backscattered microwave [34]. Accordingly, the derived NDVI, NDSI, infrared index, and RVI-based features are pointed out as: $ft^{VI} = \left[ft^{NDVI} \quad ft^{IRI} \quad ft^{NDSI} \quad ft^{RVI}\right]$. On the other hand, the statistical features, including RMS [35], Skewness [36], Kurtosis [36], and entropy [37] are explained as follows:

RMS [35]: The RMS is a type of mean. It is helpful when attempting to determine the average size of numbers where the sign is unimportant as the squaring turns all of the integers into non-negatives.

Skewness [36]: When a probability distribution differs from the symmetric normal distribution in a provided data set, it is said to be skewness *SK* [36].

Kurtosis [36]: Kurtosis, *KU*, is the degree to which outliers are present in the distribution.

Entropy [37]: The texture of the source image may be described using entropy, *EN* [37]. It is a statistical measure of unpredictability. The resulting entropy, *RMS*, Skewness, and Kurtosis properties are therefore implied as ft^{SF}; $ft^{SF} = [RMS \quad SK \quad KU \quad EN]$. The analysis of statistical feature importance is mentioned in Figure 2.

The Hydro Index (HI) is the combination of the Vegetation Index (VI) and the statistical features. The VI is represented by $\left(ft^{VI}\right)$ and the statistical feature is represented by $\left(ft^{SF}\right)$. Accordingly, Quasi Arithmetic Mean is evaluated for all statistical features (F_1) and the Weighted Geometric Mean (F_2) is evaluated for VI. Thus, NHI is computed as shown in Equation (6), in which W_1 and $W_2 \to$ weight evaluated by means of FMF as in Equation (9). Here, l points out triangular function, and a and b refers to the lower and upper limit. Here, F_1 and F_2 are evaluated as shown in Equations (7) and (8), in which w_i is computed using the chaotic cubic map.

$$NHI = (W_1 F_1 + W_2 F_2)/2 \tag{6}$$

$$F_1 = f^{-1}\left(\sum_{j=1}^{m} f(sf_j)\right) \tag{7}$$

$$F_2 = \left(\prod_{i=1}^{m} sf_j^{w_i}\right)^{1/m \sum_{i=1}^{} w_i} \tag{8}$$

$$W = \begin{cases} 0, & Y \le b \\ \frac{Y-b}{l-b}, & b \le Y \le l \\ \frac{a-Y}{a-l}, & l < Y < a \\ 0, & Y \ge a \end{cases} \tag{9}$$

The total extracted 3290 HI features are fed to the EC model namely, NN, SVM, and improved DCNN for underground water level prediction.

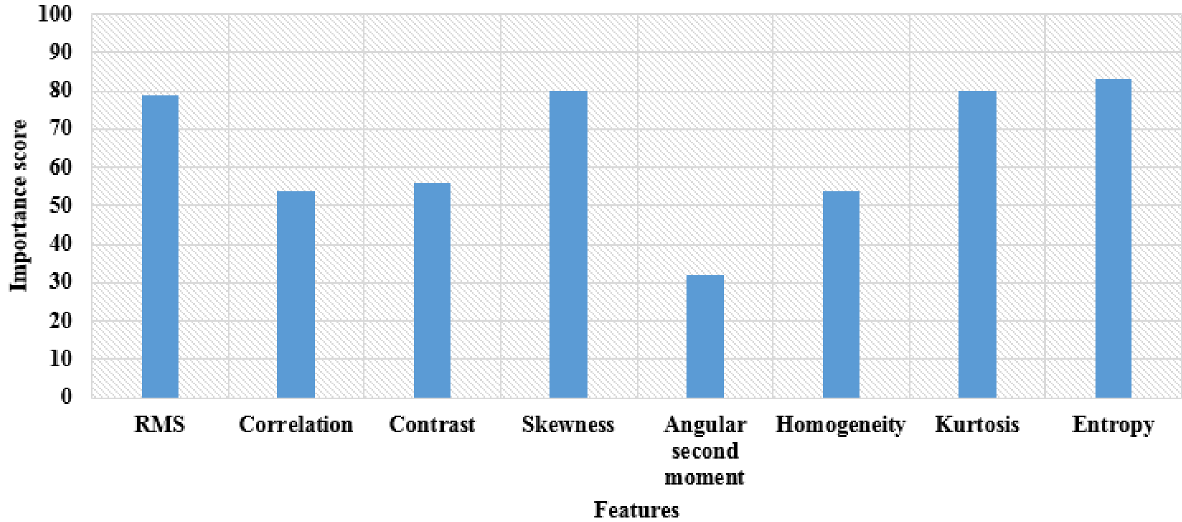


Figure 2. Analysis of statistical feature importance.

3.3. *Underground Water Level Prediction Using Ensemble Classifier*

This research works on various aspects with active and passive sensors. Active sensors can operate in a wide range of irradiation conditions but are restricted to a smaller range of wavelengths depending on the type and quantity of light sources. Unlike active sensors, which have their own light-emitting devices, passive sensors rely on sunlight as their light source. The ability of active and passive sensing systems to evaluate pertinent agronomic and physiological characteristics still needs to be understood. These types of images belong to passive remote sensing systems and those passive images are used to obtain indices and statistics. The advantage of passive sensors is that they depend on the sun's light to illuminate the target; as a result, they do not need their own energy source, making them simpler machinery. However, it is not suitable for low light levels. This information was acquired from https://earthexplorer.usgs.gov/, accessed on 1 November 2022.

The proposed ensemble model is a combination of NN, SVM, and improved DCNN. Initially, the 3290 total of extracted HI features were subjected to the NN and SVM, and the predicted results from them were subjected to improved DCNN to determine the final result.

3.3.1. Neural Network

The input passed to NN NHI that is represented in Equation (10), wherein, $m \rightarrow$ feature count in total [38].

$$HI = \{HI_1, HI_2, \ldots\ldots HI_m\} \quad (10)$$

The NN approach encompasses output, input, and hidden layers. The output $a^{(E)}$ in the hidden layer is exposed in Equation (11). The error (Er^*) should be lesser as shown in Equation (13), $s_D \rightarrow$ output neuron count, $\widehat{x}_{\hat{y}}$ and $x_{\hat{y}} \rightarrow$ predicted and actual output.

$$a^{(E)} = bf\left(w^{(V)}_{(c\hat{f})} + \sum_{e=1}^{s_f} w^{(V)}_{(\hat{e}\hat{f})} In\right) \quad (11)$$

$$\hat{x}_{\hat{y}} = bf\left(w^{(x)}_{(c\hat{y})} + \sum_{\hat{f}=1}^{s_k} w^{(x)}_{(\hat{f}\hat{y})} a^{(E)}\right) \quad (12)$$

$$Er^* = \underset{\{w^{(v)}_{(c\hat{f})}, w^{(v)}_{(\hat{e}\hat{f})}, w^{(x)}_{(c\hat{y})}, w^{(x)}_{(\hat{f}\hat{y})}\}}{\operatorname{argmin}} \sum_{y=1}^{s_D} \left| x_{\hat{y}} - \widehat{x}_{\hat{y}} \right| \quad (13)$$

In the above equations, $bf \rightarrow$ activation function, $w^{(V)}_{(c\hat{f})} \rightarrow$ Bias weight with $\hat{f}^{th}$ hidden neuron, $\hat{f}\&e \rightarrow$ neurons in hidden & input layer, $w^{(V)}_{(e\hat{f})} \rightarrow$ Weight amid e^{th} input neuron to $\hat{f}^{th}$ hidden neuron, and $\hat{s}_{\hat{k}} \rightarrow$ Input neuron count, $\hat{x}_y \rightarrow$ Network output formulated as in Equation (11). $\hat{o} \rightarrow$ Output neurons, $NH \rightarrow$ Hidden neuron count $w^{(x)}_{(c\hat{y})} \rightarrow$ Output bias weight of $\hat{y}^{th}$ output layer, $w^{(x)}_{(\hat{f}\hat{y})} \rightarrow$ Weight among $\hat{f}^{th}$ hidden layers to $\hat{y}^{th}$ the output layer. The outcome from NN is referred to as P_{nn}.

3.3.2. Support Vector Machine

SVM handles classification problems and simpler nonlinear regression issues [39]. Lagrange's work has two problems that should be taken into account, as they make quadratic programming problems less problematic in the SVM model.

$$\begin{aligned} -L(\kappa) = -\sum_{d=1}^{Q} \gamma_d + \\ \tfrac{1}{2} \cdot \sum_{d=1}^{J} \sum_{h=1}^{J} \gamma_d \cdot \gamma_h . X_d . X_h . \kappa(z_d,\ z_h) \quad &\rightarrow \min_x, \\ \sum_{d=1}^{J} \gamma_d, X_d = 0,\ 0 \leq \gamma \leq Cr,\ d = \overline{1, J} \end{aligned} \quad (14)$$

In Equation (14), $z_d \rightarrow$ data presented throughout training, $\gamma_d \rightarrow$ dual term, $Y_d \rightarrow +1$ or $-1$and $\kappa(z_d, z_h), \rightarrow$ kernels function. Furthermore, $J \rightarrow$ data count in training dataset and $Cr \rightarrow$ regularization factor. The outcome from SVM is referred to as P_{svm}. The SVM and NN outputs are fused and given to improved DCNN for final prediction.

3.3.3. Improved Deep Convolutional Neural Network

The well known DL model is DCNN [40]. The convolution layer often includes some convolution kernels that help with the calculation of different feature maps. Particularly, each neuron in the feature map is connected to a neuron in the layer above. The feature values are evaluated as in Equation (15) in position $(a\ ,\ b)$ at lth layer of the matching cth feature map. The bias and weight of cth the filter are referred as B^l_c and W^l_c, $Y^l_{a,b}$ refers to patched input.

$$Z^l_{a,b,c} = W^{l\,T}_c Y^l_{a,b} + B^l_c \quad (15)$$

The activation value $\left(act^{l}_{a,b,c}\right)$ linked to the convolutional feature $Z^{l}_{a,b,c}$ is evaluated as in Equation (16).

$$act^{l}_{a,b,c} = act\left(Z^{l}_{a,b,c}\right) \tag{16}$$

Utilizing the convolutional layer outputs, the pooling layers of CNN perform down-sampling operations. According to Equation (17), each feature map at the present layer is related to the following feature map at the pooling layer.

$$V^{l}_{a,b,c} = pool\left(act^{l}_{m,n,c}\right), \forall(m.n) \in R_{a,b} \tag{17}$$

As a newly improved DCNN, the reconstruction error is modeled as in Equation (18), in which u refers to the number of training samples, v refers to features in each group of samples, P_{ab} refers to the reconstructed value of RBM training samples, y_{ab} refers to the count of values.

$$RC_{error-new} = \frac{\sum_{i=1}^{u}\sum_{i=1}^{v} P_{ab} - y_{ab}}{uvP_y} \tag{18}$$

$$Loss = \frac{RC_{error} + CE}{2} \tag{19}$$

Thus, loss function is modeled as in Equation (19). The prediction result will come from the improved DCNN model that predicts the water level as Critical, Overexploited, Safe, Saline, or Semi-critical. The assumed parameters of the EC are: activation function is ReLU, learning rate is 0.0001, loss function is cross entropy, Kernel is radial basis function, number of epochs is 100, and optimizer is Adam. The results and discussion of the proposed framework are mentioned in Section 4.

4. Results and Discussion

4.1. Simulation Procedure

The developed underground water level prediction framework using remote sensing image is simulated using the MATLAB 2021b version 9.11 with a RAM of 16GB and Intel Core i3 12th generation processor. Despite many works available on underground water level prediction, only few research works fall under DL strategy to solve the given prediction process, hence, we are validating the performance of the proposed work over standard DL models, such as (LSTM, NB, RF, RNN, Bi-GRU, ML) [20], DL [21], CNN [22], EC [28], and TS + RF [41]. The dataset was collected from https://earthexplorer.usgs.gov/, accessed on 1 November 2022. Here, the image is divided into five categories: semi-critical, safe, overexploited, safe, and critical. Each category of the image is classified into one of five types: critical, overexploited, safe, saline, and semi-critical. The underground water level prediction has been performed by categorizing the levels of underground water into five levels as explained: (i) Critical: Indicates that the level of water is less, (ii) Overexploited: Indicates the over-consumed water level, (iii) Safe: Indicates that the level of water is sufficient, (iv) Saline: Indicates the salinity level of water, and (v) Semi-critical: Indicates the semi-critical level of water.

$$F - measure = \frac{2TP}{2TP + FP + FN} \tag{20}$$

$$FPR = \frac{FP}{FP + TN} \tag{21}$$

$$Specificity = \frac{TN}{TN + FP} \tag{22}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{23}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{24}$$

The performance measures: F-measure, False Positive Rate (FPR), specificity, precision, accuracy, Matthews Correlation Coefficient (MCC), Negative Predictive Values (NPV), sensitivity, and False Negative Rate (FNR) are utilized for validating the proposed framework's effectiveness, which are mathematically specified in Equations (20)–(28). True Positives, True Negatives, False Positives, and False Negatives are stated as TP, TN, FP, and FN.

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}} \tag{25}$$

$$NPV = \frac{TN}{TN + FN} \tag{26}$$

$$Sensitivity = \frac{TP}{FN + TP} \tag{27}$$

$$FNR = \frac{FN}{FN + TP} \tag{28}$$

4.2. Location Specification

Saline

Andhra Pradesh-Nagayalanka: 15.9455 latitude, 80.9180 longitude—Water source level range—77 sq.km.
Tamil Nadu-Nagapattinam: 10.7672 latitude, 79.8449 longitude—Water source level range—27.83 sq.km.

Semi critical

Andhra Pradesh-Kadapa-Ramapuram: 14.8080 latitude, 78.7072 longitude—Water source level range—79 sq.km.
Madhya Pradesh-Manasa: 24.4748 latitude, 75.1404 longitude—Water source level range—48 sq.km.

Safe

Telangana-Mahabubnagar-Maddur: 16.8602 latitude, 77.6121 longitude—Water source level range—184 sq.km.
Andhra Pradesh-Visakhapatnam-Kotapadu: 17.8861 latitude, 83.0435 longitude—Water source level range—352 sq.km.

Exploited

Rajasthan-Jalor-Jaswantpura: 24.8019 latitude, 72.4598 longitude—Water source level range—64 sq.km.
Rajasthan-Nagaur: 27.1983 latitude, 73.7493 longitude—Water source level range—77 sq.km.

Critical

Karnataka-Raichur: 16.2160 latitude, 77.3566 longitude—Water source level range—83 sq.km.
Rajasthan-Bharatpur: 27.2152 latitude, 77.5030 longitude—Water source level range—44.10 sq.km.

Figure 3 shows images of regions where the underground water level is in critical condition, as well as the preprocessed images. Figure 4 explains the images of regions where the underground water level is highly exploited i.e., overconsumed, as well as the corresponding preprocessed images. Figure 5 shows the images of regions with safe water level and the preprocessed images of two sample images. Figure 6 shows the images of the regions with saline water and the preprocessed image results using Wiener filtering. Figure 7 shows the images of the regions under the semi-critical level and the preprocessed images as well.

Figures 3 and 4 show the representation of underground water level images at critical and highly exploited conditions, where, the pixel resolution level is poor and unable to gather more information from the particular image. Figure 5 depicts images of regions processed as those of safe water level, where we can get huge volumes of information with higher accuracy from the particular image. Figure 6 shows the representation of underground water level images with salinity conditions. Due to the salinity process, we are unable to gain clear portions from the image. Figure 7 shows the representation of underground water level images at semi-critical conditions, where we can get average data with acceptable accuracy. The location and its evaluation ranges of all the specified images are already declared in the initial stage.

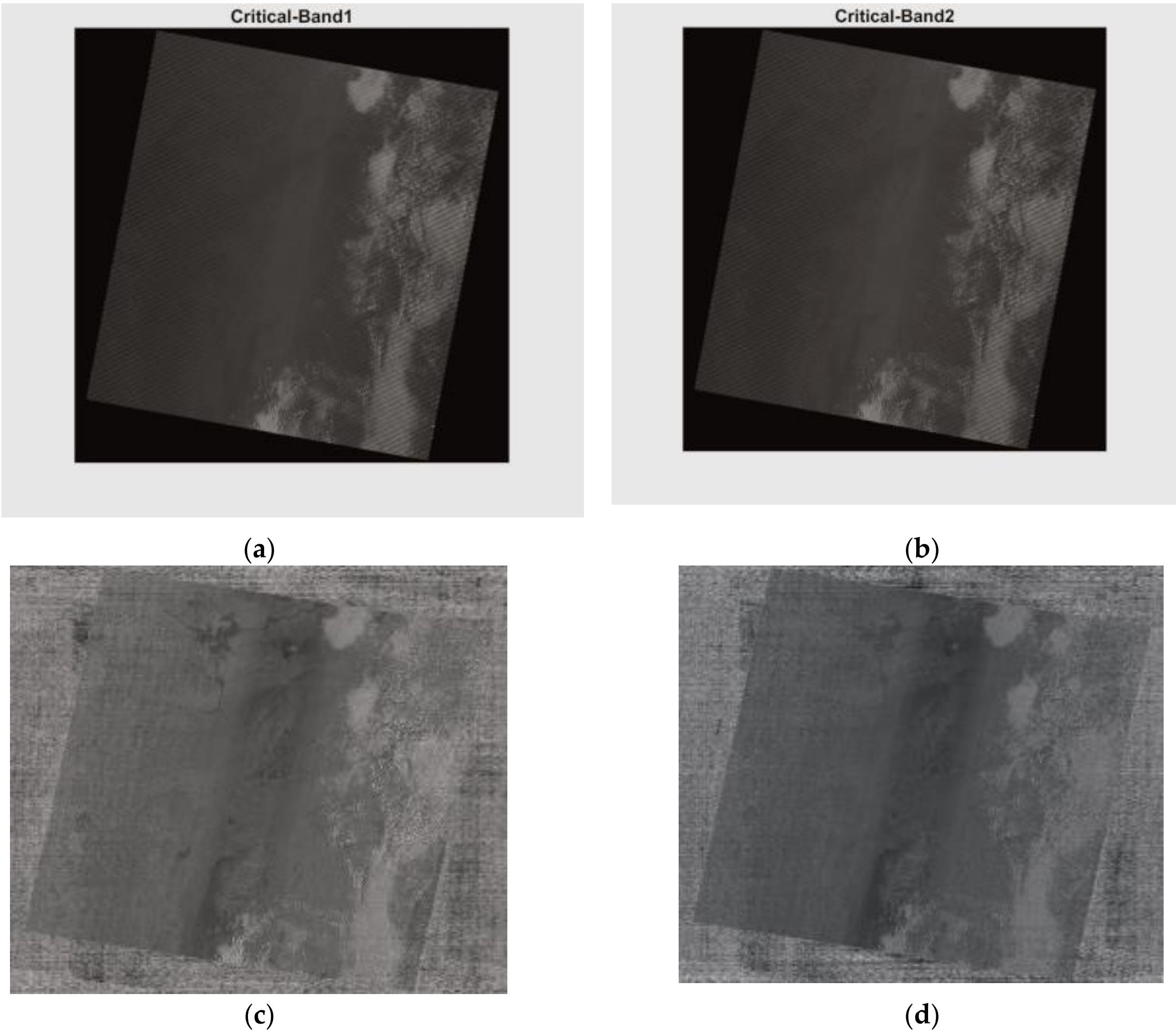
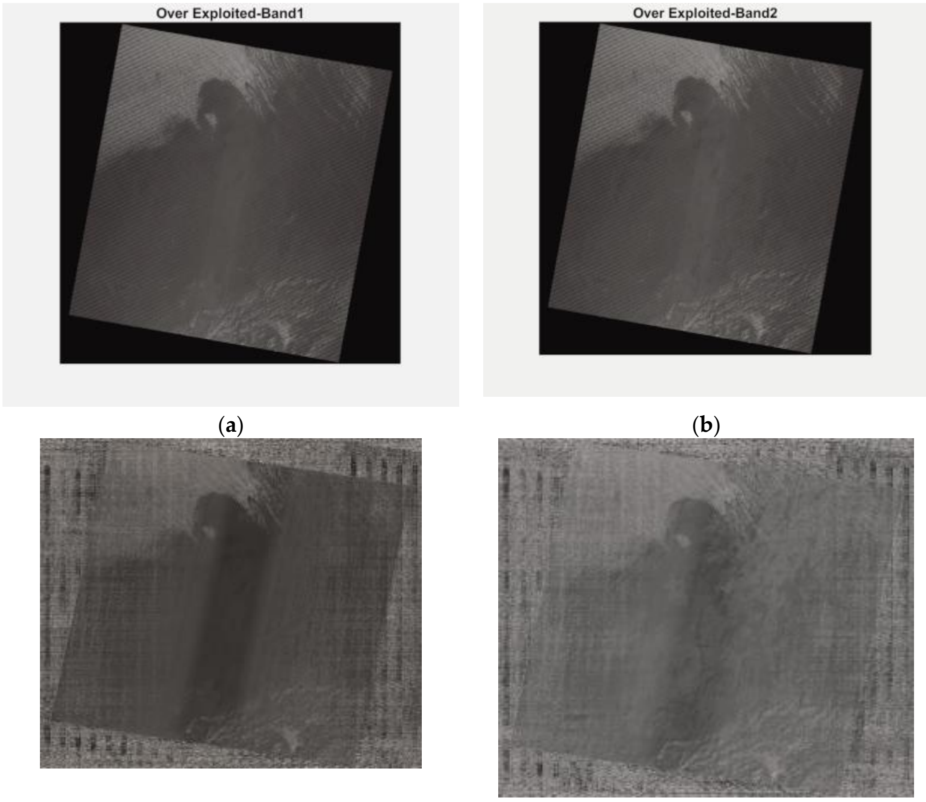
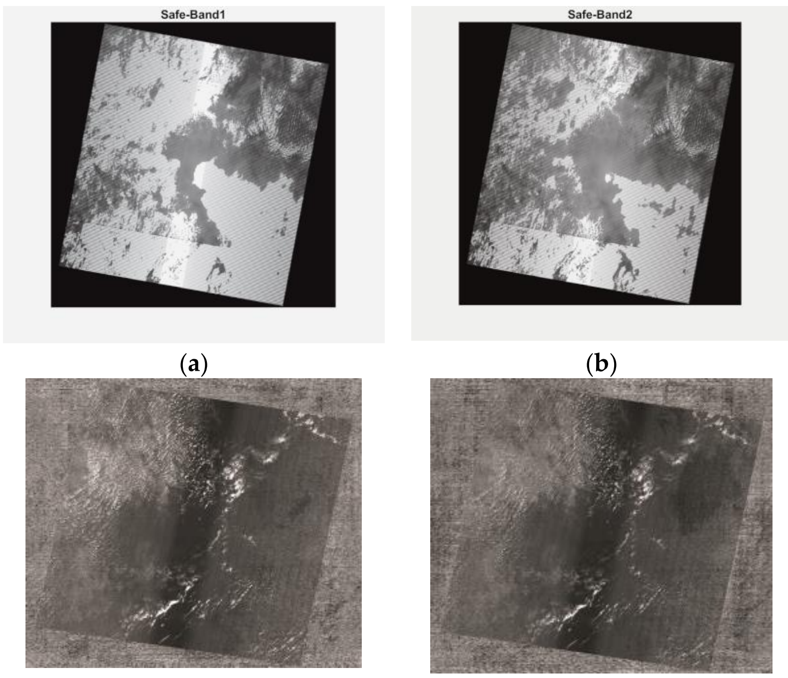


Figure 3. Images of regions where the underground water level is in critical condition. (**a**,**b**) Images of two different regions and (**c**,**d**) preprocessed image using Wiener filtering.



(**a**) (**b**) (**c**) (**d**)

Figure 4. Images of regions where the underground water level is highly exploited i.e., overconsumed. (**a**) Sample image 1 and (**b**) sample image 2. (**c**) Preprocessed image 1 by using Wiener filtering. (**d**) Preprocessed image 2 by using Wiener filtering.



(**a**) (**b**) (**c**) (**d**)

Figure 5. Images of regions with Safe water level. (**a**) Original image 1 and (**b**) original image 2. (**c**) Preprocessed image 1 by using Wiener filtering. (**d**) Preprocessed image 2 by using Wiener filtering.

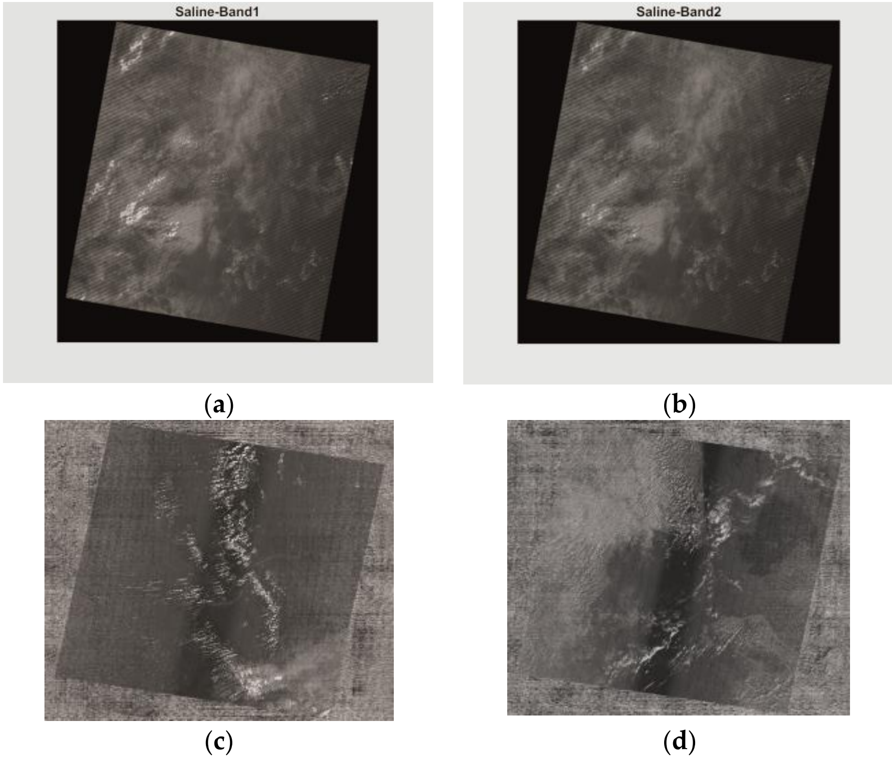


Figure 6. Images of regions with salinity (saline water). (**a**) Original image 1 and (**b**) original image 2. (**c**) Preprocessed image 1 by using Wiener filtering. (**d**) Preprocessed image 2 by using Wiener filtering.

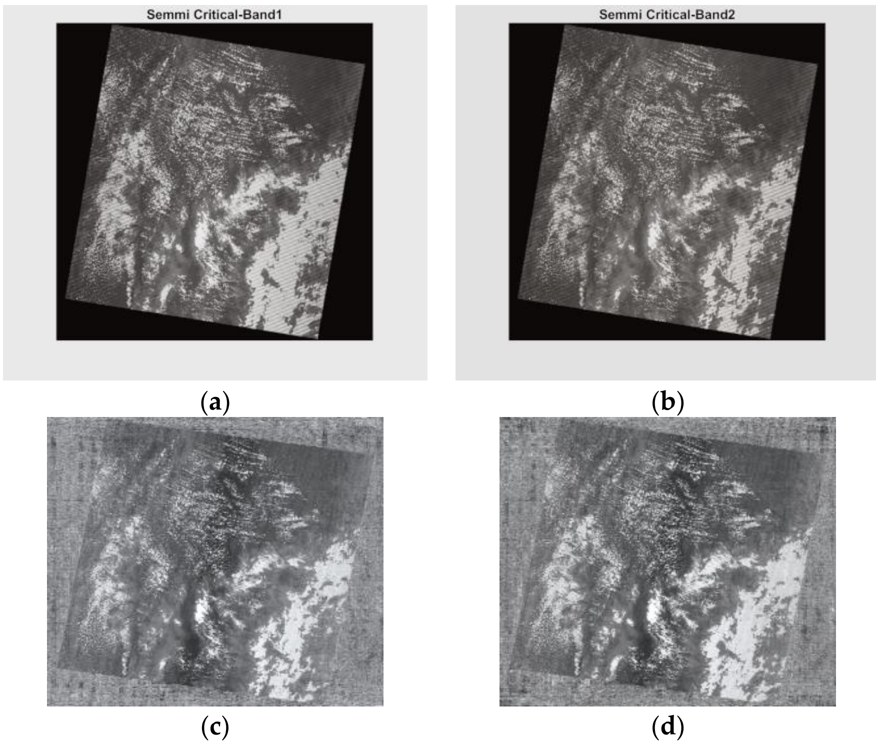


Figure 7. Images of regions with Semi-critical. (**a**) Original image 1 and (**b**) original image 2. (**c**) Preprocessed image 1 by using Wiener filtering. (**d**) Preprocessed image 2 by using Wiener filtering.

4.3. Performance Analysis

In this work, we have considered 84 images for the Safe category, 35 images for Semi-critical, 14 Critical images, 14 images for the Saline category, and 28 images for the Overexploited category. In total, 175 images are considered. This section shows the performance analysis by varying the training percentage. For example, for 70% of the training rate, the number of training images is 123 and the number of testing images is 52.

At the 90th LP, the specificity of the adopted EC approach is significantly greater when compared with other LPs. The EC model has greater sensitivity and precision at the 80th and 90th LPs than at the 60th and 70th. At the 70th LP, the F-measure using the EC scheme is higher, whereas at the 60th LP, the F-measure using the EC scheme is higher. While the FNR is nearly similar at other LPs, it is much less at the 70th LP using the EC scheme. While the FPR at other LPs using the EC scheme is nearly similar, it is much higher at the 70th LP using the EC scheme. The MCC using the EC scheme has gained a high value of 0.99 at the 90th LP. The NPV using the EC scheme is high at 70th LP, with the next high value of NPV using the EC scheme acquired at the 90th LP with a value of 0.9.

In addition, the analysis of the EC scheme over extant classifiers, such as LSTM, NB, RF, RNN, Bi-GRU, EC [28], and TS + RF [41] is shown in Table 1. The accuracy in Table 1 is high at the 90th LP with 0.9%, while LSTM, NB, RF, RNN, Bi-GRU, EC [28], ML [20], DL [21], CNN [22], and TS + RF [41] have acquired relatively lesser values of 0.881167, 0.898583, 0.917667, 0.891417, 0.878889, 0.892167, 0.8765, 0.8654, 0.8432, and 0.898583 in that order. This observation proves the performance of the proposed work on predicting water levels under the given categories.

Table 1. Performance study on EC over extant classifiers in predicting the category of underground water levels and their conditions.

	LSTM	NB	RF	RNN	BI-GRU	EC [28]	ML [20]	DL [21]	CNN [22]	TS + RF [41]	EC (NN, SVM, DCNN)
F-measure	0.915	0.927	0.940	0.908	0.914	0.909	0.717	0.846	0.623	0.927	0.957
FPR	0.332	0.283	0.229	0.003	0.337	0	0.435	0.524	0.576	0.283	0.008
Specificity	0.667	0.716	0.770	0.996	0.662	1	0.717	0.762	0.788	0.716	0.929
Precision	0.845	0.865	0.888	0.997	0.843	1	0.435	0.524	0.576	0.865	0.920
Accuracy	0.881	0.898	0.917	0.891	0.878	0.892	0.435	0.524	0.576	0.898	0.928
MCC	0.747	0.784	0.824	0.796	0.742	0.798	0.152	0.286	0.364	0.784	0.928
NPV	0.994	0.995	0.995	0.766	0.992	0.766	0.717	0.762	0.788	0.995	1
Sensitivity	0.998	0.998	0.998	0.834	0.997	0.833	0.282	0.237	0.211	0.998	0.925
FNR	0.001	0.001	0.001	0.165	0.002	0.166	0.564	0.475	0.423	0.001	0

4.4. Statistical Analysis

Table 2 shows the error statistics of the deployed EC model over LSTM, NB, RF, RNN, ML [20], DL [21], CNN [22], and Bi-GRU in categorizing the underground water levels and their conditions from 10-fold experimental validation. With this method, we have a data set that we divided randomly into 10 parts. We used nine of those parts for training and one tenth of them for testing. We repeated this procedure 10 times, each time reserving a different tenth for testing. Based on this, the statistical analysis was carried out with respect to error deviation, minimum error, mean error, maximum error, and error median. From Table 2, the deployed EC system demonstrated enhanced outcomes over LSTM, NB, RF, RNN, ML [20], DL [21], CNN [22], and Bi-GRU schemes for all statistical cases. Predominantly, the deployed EC method offers less error over LSTM, NB, RF, RNN, and Bi-GRU. This shows that the accomplishments of compared classifiers were not as superior as EC, whereas the EC model makes the prediction more precise. Particularly, the obtained mean error rate is less than 0.1 for the proposed EC model, whereas the rest of the conventional methods give a high error rate. Additionally, the maximum error rate obtained by the proposed EC model is 0.080, which is not the case of other models, as they show more than 0.1. Therefore, the enhancement of the EC model is proven over others with derived features and improved DCNN. The overall error statistics by means of error deviation, minimum error, mean error, maximum error, and error median states that the proposed EC model achieved superior results in underground water level prediction with a minimum error value, whereas the overall error statistics indicate that the proposed EC model is able to achieve higher results on other complex real-time datasets with a low error value.

Table 2. Error statistics achieved by EC and existing classifiers in categorizing underground water levels and their conditions from 10-fold experimental validation.

Models	Error Deviation	Minimum Error	Mean Error	Maximum Error	Error Median
EC	0.008	0.059	0.067	0.080	0.065
LSTM	0.033	0.090	0.114	0.172	0.098
NB	0.023	0.092	0.121	0.156	0.119
RF	0.027	0.134	0.157	0.201	0.148
RNN	0.036	0.163	0.127	0.231	0.140
Bi-GRU	0.025	0.102	0.133	0.162	0.133
ML [20]	0.122	0.187	0.160	0.165	0.023
DL [21]	0.119	0.153	0.134	0.132	0.012
CNN [22]	0.124	0.187	0.151	0.147	0.022

4.5. Analysis of Features on Statistical Errors

The analysis of the adopted EC scheme overrepresented the method without statistical features, the presented method without a novel hydro index, and the presented method without features as shown in Table 3. In Table 3, the NPV of EC is higher than the presented method without statistical features, the presented method without novel hydro index, and the presented method without features. Next to the EC scheme, the presented method without features has achieved superior NPV outcomes over the adopted method without statistical features and the presented method without a novel hydro index. Likewise, in Table 3, the FPR of EC is lesser (0.008) than the presented method without statistical features, the presented method without a novel hydro index, and the presented method without features. Next to the EC scheme, the presented method without features has achieved fewer FPR outcomes over the adopted method without statistical features and the presented method without a novel hydro index. Similarly, the F-measure of EC is high (0.957) over the presented method without statistical features, the presented method without a novel hydro index, and the presented method without features. Next to the EC scheme, the presented method without statistical features has achieved superior F-measure (0.930) outcomes over the adopted method without features and the presented method without a novel hydro index.

Table 3. Error statistics achieved by EC, which is trained by different features, in categorizing underground water levels and their conditions from 10-fold experimental validation.

	EC without Statistical Feature	EC without Novel Hydro Index	EC without Feature Extraction	Proposed EC
MCC	0.009	0.013	0.003	0.928
NPV	0.073	0.100	0.146	0.851
Specificity	0.046	0.044	0.051	0.929
FNR	0.053	0.054	0.053	0.021
F-measure	0.930	0.911	0.894	0.957
Accuracy	0.870	0.838	0.810	0.928
Sensitivity	0.946	0.945	0.946	0.925
FPR	0.953	0.955	0.948	0.008
Precision	0.915	0.880	0.847	0.920

4.6. Discussion

One of the most vital natural elements is groundwater. In order to comprehend groundwater resources and maintain sustainable development of water resources and the environment, regional groundwater depth is a crucial parameter. Here, the Improved Hydro Index Value with Ensemble Classifier was used to evaluate the prediction of underground water levels in remote sensing images. A Wiener filter was developed for removing additive noise and blurring from obtained remote sensing images. The Wiener filter is best because it minimizes the total mean square error during the inverse filtering and noise smoothing processes. Statistical features and integrated VI, NDVI, NDSI, IRI, and RVI features were used for feature extraction. The effectiveness of predicting underground water levels using remote sensing images is improved by the discriminative extracted features, which reduce the sematic space between the feature subsets. Improved NN, SVM, and DCNN are included in the Proposed Ensemble Classifier (EC) model for accurate underground water level forecast. According to the findings, the suggested EC model's False Positive Rate (FPR) is 0.0083, which is lower than that of current approaches. The proposed EC model, however, outperforms conventional models, including the Long Short-Term Memory (LSTM) network, Naive Bayes (NB), Random Forest (RF), Recurrent Neural Network (RNN), and Bidirectional Gated Recurrent Unit (Bi-GRU) with a high accuracy of 0.90.

5. Conclusions

This research suggested a new framework for predicting the underground level of water using remote sensing images. Firstly, a Wiener filter is used for preprocessing, and further, features including VI (NDVI, NDSI, IRI, and RVI) and statistical features (entropy, RMS, Skewness, and Kurtosis) are extracted from preprocessed images. Here, the VI and statistical features are combined as a novel hydro index. Finally, the extracted 3290 HI features are given to the EC model (NN, SVM, and improved DCNN) for underground water level prediction. According to the resulting findings, the proposed EC model's False Positive Rate (FPR) is 0.0083, which is lower than that of conventional approaches. The proposed EC model, in contrast, has a high accuracy of 0.90, which is better than the conventional models. Further, the MCC of the EC model has gained a high value of 0.99 at the 90th LP, the NPV of the EC model is high at the 70th LP, and the next high NPV value of the EC model is obtained at the 90th LP with a value of 0.90. However, utilizing DL models raises many issues: DL models are frequently interpreted as black boxes that imply that it is very challenging to unravel and recognize the automatic feature selection process, which finally occurs, and the forecasts that emerge from any given DL-oriented schemes. They are also very expensive to train and utilize, regarding time and resources. In the future, this research will be further extended by analyzing various simulation settings with novel hybrid methods for feature selection.

Author Contributions: Paper investigation, resources, data curation, writing—original draft preparation, writing—review and editing, and visualization were conducted by S.C.N. and P.V. The paper conceptualization and software was conducted by N.S.G. and K.S. The validation and formal analysis, methodology, supervision, project administration, and funding acquisition of the version to be published were conducted by A.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Nomenclature

Abbreviation	Description
Bi-GRU	Bidirectional Gated Recurrent Unit
DCNN	Deep Convolutional Neural Network
DL	Deep Learning
EC	Ensemble Classifier
DBN	Deep Belief Network
GIS	Geographic Information Systems
GWL	Ground Water Level
ConvLSTM	Convolutional Long-Short-Term Memory
FMF	Fuzzy Membership Function
EBF	Evidential Belief Function
HRSI	Hyper Spectral Remote Sensing Image
LSTM	Long Short-Term Memory
ROI	Region Of Interest
LP	Learning Percentage
ML	Machine Learning
MMSE	Minimal Mean Square Error
SWIR	Shorter Wave Infra Red Band
NIR	Normalized Infra Red Ratio
NN	Neural Network
NB	Naïve Bayes
LR	Logistic Regression
NDVI	Normalized Difference Vegetation Index
NDSI	Normalized Difference Snow Index
RVI	Radar Vegetation Index
IRI	Infra Red Index
RNN	Recurrent Neural Network
RMS	Root Mean Square
RF	Random Forest
TS-RF	Temporal Segmentation-Random Forest
SVM	Support Vector Machine
VI	Vegetation Index

References

1. Fathi, A.; Lee, T.; Mohebzadeh, H. Allocating Underground Dam Sites Using Remote Sensing and GIS Case Study on the Southwestern Plain of Tehran Province, Iran. *J. Indian Soc. Remote Sens.* **2019**, *47*, 989–1002. [CrossRef]
2. Yan, S.; Shi, K.; Li, Y.; Liu, J.; Zhao, H. Integration of Satellite Remote Sensing Data in Underground Coal Fire Detection: A Case Study of the Fukang Region, Xinjiang, China. *Front. Earth Sci.* **2020**, *14*, 1–12. [CrossRef]
3. Wang, X.; Guo, Z.; Wu, L.; Zhu, C.; He, H. Extraction of Palaeochannel Information from Remote Sensing Imagery in the East of Chaohu Lake, China. *Front. Earth Sci.* **2012**, *6*, 75–82. [CrossRef]
4. Fossi, D.H.; Djomo, H.D.; Takodjou Wambo, J.D.; Kouayep Tchoundi, L.C.; Deassou Sezine, E.; Takam Tchoupe, G.B.; Tchatchueng, R. Extraction and Analysis of Structural Lineaments from Mokolo Area, North Cameroon, Using DEM and Remote Sensing Images, and Their Influence on Drainage Morphometric. *Arab. J. Geosci.* **2021**, *14*, 2062. [CrossRef]
5. Guo, F.; Gao, Z. RETRACTED ARTICLE: Sponge City Plant Planning and Urban Construction Based on High-Resolution Remote Sensing Images. *Arab. J. Geosci.* **2021**, *14*, 1131. [CrossRef]
6. İrdemez, Ş.; Eymirli, E.B. Determination of Spatiotemporal Changes in Erzurum Plain Wetland System Using Remote Sensing Techniques. *Environ. Monit. Assess.* **2021**, *193*, 265. [CrossRef]
7. Siming, C.; Aidi, H.; Wenke, G. Remote Sensing Monitoring Method for Groundwater Level on Aeolian Desertification Area. *J. Water Chem. Technol.* **2020**, *42*, 522–529. [CrossRef]
8. Zacharias, I.; Dimitriou, E.; Koussouris, T. Quantifying Land-Use Alterations and Associated Hydrologic Impacts at a Wetland Area by Using Remote Sensing and Modeling Techniques. *Environ. Model. Assess.* **2004**, *9*, 23–32. [CrossRef]
9. Liu, H.; Jiang, Y.; Misa, R.; Gao, J.; Xia, M.; Preusse, A.; Sroka, A.; Jiang, Y. Ecological Environment Changes of Mining Areas around Nansi Lake with Remote Sensing Monitoring. *Environ. Sci. Pollut. Res.* **2021**, *28*, 44152–44164. [CrossRef]
10. Jha, M.K.; Chowdhury, A.; Chowdary, V.M.; Peiffer, S. Groundwater Management and Development by Integrated Remote Sensing and Geographic Information Systems: Prospects and Constraints. *Water Resour. Manag.* **2007**, *21*, 427–467. [CrossRef]

11. Joshi, P.K.; Kumar, M.; Midha, N.; Yanand, V.; Wal, A.P. Assessing Areas Deforested by Coal Mining Activities through Satellite Remote Sensing Images and Gis in Parts of Korba, Chattisgarh. *J. Ind. Soc. Remote Sens.* **2006**, *34*, 415–421. [CrossRef]
12. Sivasankar, T.; Borah, S.B.; Das, R.; Raju, P.L.N. An Investigation on Sudden Change in Water Quality of Brahmaputra River Using Remote Sensing and GIS. *Natl. Acad. Sci. Lett.* **2020**, *43*, 619–623. [CrossRef]
13. Lee, S.; Hyun, Y.; Lee, S.; Lee, M.J. Groundwater potential mapping using remote sensing and GIS-based machine learning techniques. *Remote Sens.* **2020**, *12*, 1200. [CrossRef]
14. Collignon, B. A New Tool for the Remote Sensing of Groundwater Tables: Satellite Images of Pastoral Wells. *Open Geospat. Data Softw. Stand.* **2020**, *5*, 4. [CrossRef]
15. Chowdary, V.M.; Ramakrishnan, D.; Srivastava, Y.K.; Chandran, V.; Jeyaram, A. Integrated Water Resource Development Plan for Sustainable Management of Mayurakshi Watershed, India Using Remote Sensing and GIS. *Water Resour. Manag.* **2009**, *23*, 1581–1602. [CrossRef]
16. Molla, M.H.; Chowdhury, M.A.T.; Islam, A.M.Z. Spatiotemporal Change of Urban Water Bodies in Bangladesh: A Case Study of Chittagong Metropolitan City Using Remote Sensing (RS) and GIS Analytic Techniques. *J. Ind. Soc. Remote Sens.* **2021**, *49*, 773–792. [CrossRef]
17. Cheng, C.; Zhang, F.; Shi, J.; Kung, H.-T. What Is the Relationship between Land Use and Surface Water Quality? A Review and Prospects from Remote Sensing Perspective. *Environ. Sci. Pollut. Res.* **2022**, *29*, 56887–56907. [CrossRef]
18. El Garouani, A.; Aharik, K.; El Garouani, S. Water Balance Assessment Using Remote Sensing, Wet-Spass Model, CN-SCS, and GIS for Water Resources Management in Saïss Plain (Morocco). *Arab. J. Geosci.* **2020**, *13*, 738. [CrossRef]
19. Jiang, H.; Liu, C.; Sun, X.; Lu, J.; Zou, C.; Hou, Y.; Lu, X. Remote Sensing Reversion of Water Depths and Water Management for the Stopover Site of Siberian Cranes at Momoge, China. *Wetlands* **2015**, *35*, 369–379. [CrossRef]
20. Majumdar, S.; Smith, R.; Butler, J.J.; Lakshmi, V. Groundwater Withdrawal Prediction Using Integrated Multitemporal Remote Sensing Data Sets and Machine Learning. *Water Resour. Res.* **2020**, *56*, e2020WR028059. [CrossRef]
21. Sureshkumar, V.; Somarajadikshitar, R.; Beeram, B.S. A Novel Representation and Prediction Initiative for Underground Water by Using Deep Learning Technique of Remote Sensing Images. *Comput. J.* **2022**, bxac101. [CrossRef]
22. Wang, Z.; Tian, S. Ground Object Information Extraction from Hyperspectral Remote Sensing Images Using Deep Learning Algorithm. *Microprocess. Microsyst.* **2021**, *87*, 104394. [CrossRef]
23. Suganthi, S.; Elango, L.; Subramanian, S.K. Groundwater potential zonation by Remote Sensing and GIS techniques and its relation to the Groundwater level in the Coastal part of the Arani and Koratalai River Basin, Southern India. *Earth Sci. Res. J.* **2013**, *17*, 87–95.
24. Tao, H.; Hameed, M.M.; Marhoon, H.A.; Zounemat-Kermani, M.; Heddam, S.; Kim, S.; Sulaiman, S.O.; Tan, M.L.; Sa'adi, Z.; Mehr, A.D.; et al. Groundwater Level Prediction Using Machine Learning Models: A Comprehensive Review. *Neurocomputing* **2022**, *489*, 271–308. [CrossRef]
25. Shnewer, F.M.; Hasan, A.A.; AL-Zuhairy, M.S. Groundwater Site Prediction Using Remote Sensing, GIS and Statistical Approaches: A Case Study in the Western Desert, Iraq. *IJET* **2018**, *7*, 166–173. [CrossRef]
26. Zipper, S.C.; Gleeson, T.; Kerr, B.; Howard, J.K.; Rohde, M.M.; Carah, J.; Zimmerman, J. Rapid and accurate estimates of streamflow depletion caused by groundwater pumping using analytical depletion functions. *Water Resources Research* **2019**, *55*, 5807–5829. [CrossRef]
27. El-Hadidy, S.M.; Morsy, S.M. Expected Spatio-Temporal Variation of Groundwater Deficit by Integrating Groundwater Modeling, Remote Sensing, and GIS Techniques. *Egypt. J. Remote Sens. Space Sci.* **2022**, *25*, 97–111. [CrossRef]
28. Hussein, E.A.; Thron, C.; Ghaziasgar, M.; Bagula, A.; Vaccari, M. Groundwater Prediction Using Machine-Learning Tools. *Algorithms* **2020**, *13*, 300. [CrossRef]
29. Zhang, H.; Ma, J.; Chen, C.; Tian, X. NDVI-Net: A fusion network for generating high-resolution normalized difference vegetation index in remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2020**, *168*, 182–196. [CrossRef]
30. Gascoin, S.; Dumont, Z.B.; Deschamps-Berger, C.; Marti, F.; Salgues, G.; López-Moreno, J.I.; Revuelto, J.; Michon, T.; Schattan, P.; Hagolle, O. Estimating fractional snow cover in open terrain from sentinel-2 using the normalized difference snow index. *Remote Sens.* **2020**, *12*, 2904. [CrossRef]
31. Gonenc, A.; Ozerdem, M.S.; Acar, E. Comparison of NDVI and RVI Vegetation Indices Using Satellite Images. In Proceedings of the 2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics), Istanbul, Turkey, 16–19 July 2019; pp. 1–4. [CrossRef]
32. Kim, Y.; van Zyl, J.J. A Time-Series Approach to Estimate Soil Moisture Using Polarimetric Radar Data. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 2519–2527. [CrossRef]
33. Koppe, W.; Gnyp, M.L.; Hütt, C.; Yao, Y.; Miao, Y.; Chen, X.; Bareth, G. Rice Monitoring with Multi-Temporal and Dual-Polarimetric TerraSAR-X Data. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *21*, 568–576. [CrossRef]
34. Wen, Q.; Liu, H.; Zhang, Z. A method of generating multivariate non-normal random numbers with desired multivariate skewness and kurtosis. *Behav. Res. Methods* **2020**, *52*, 939–946.
35. Martin, C.; Aleem, S.H.A.; Zobaa, A.F. On the root mean square error (RMSE) calculation for parameter estimation of photovoltaic models: A novel exact analytical solution based on Lambert W function. *Energy Convers. Manag.* **2020**, *210*, 112716.
36. Gawali, S. Shape of Data: Skewness and Kurtosis. 2021. Available online: https://www.analyticsvidhya.com/blog/2021/05/shape-of-data-skewness-and-kurtosis/ (accessed on 5 September 2022).

37. Yan, H.; Deng, H. An improved belief entropy in evidence theory. *IEEE Access* **2020**, *8*, 57505–57516. [CrossRef]
38. Mohan, Y.; Chee, S.S.; Xin, D.K.P.; Foong, L.P. Artificial Neural Network for Classification of Depressive and Normal in EEG. In Proceedings of the 2016 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES), Kuala Lumpur, Malaysia, 4–8 December 2016; pp. 286–290. [CrossRef]
39. Avci, E. A New Intelligent Diagnosis System for the Heart Valve Diseases by Using Genetic-SVM Classifier. *Expert Syst. Appl.* **2009**, *36*, 10618–10626. [CrossRef]
40. Gu, J.; Wang, Z.; Kuen, J.; Ma, L.; Shahroudy, A.; Shuai, B.; Liu, T.; Wang, X.; Wang, G.; Cai, J.; et al. Recent Advances in Convolutional Neural Networks. *Pattern Recognit.* **2018**, *77*, 354–377. [CrossRef]
41. Rohde, M.M.; Biswas, T.; Housman, I.W.; Campbell, L.S.; Klausmeyer, K.R.; Howard, J.K. A Machine Learning Approach to Predict Groundwater Levels in California Reveals Ecosystems at Risk. *Front. Earth Sci.* **2021**, *9*, 784499. [CrossRef]

Multimedia Tools and Applications (2023) 82:20989–21004
https://doi.org/10.1007/s11042-023-14460-8

A proficient video recommendation framework using hybrid fuzzy C means clustering and Kullback-Leibler divergence algorithms

H. Anwar Basha[1] • **S. K. B Sangeetha**[2] • **S. Sasikumar**[3] • **J. Arunnehru**[2] • **M. Subramaniam**[4]

Received: 14 April 2021 / Revised: 9 February 2022 / Accepted: 31 January 2023 /
Published online: 7 February 2023

Abstract
A video recommendation framework for e-commerce clients is proposed using the collaborative filtering (CF) process. One of the most important features of the CF algorithm is its scalability. To avoid the issue, a hybrid model-based collaborative filtering approach is proposed. KL Divergence was developed to address the CF technique's scalability problem. The clustering with enhanced sqrt-cosine similarity Recommender scheme is proposed. For successful clustering, Kullback–Leibler Divergence-based Fuzzy C-Means clustering is suggested, with the aim of focusing on greater accuracy during movie recommendation.The proposed scheme is viewed as a trustworthy contribution that significantly improves the ability of movie recommendation by virtue of the KL divergence-based Fuzzy C-Means clustering mechanism and enhanced sqrt-

✉ H. Anwar Basha
anwar.mtech@gmail.com

S. K. B Sangeetha
sangeets8@srmist.edu.in

S. Sasikumar
sasikumar@saveetha.ac.in

J. Arunnehru
arunnehru.aucse@gmail.com

M. Subramaniam
subbu.21074@gmail.com

1 School of Computer Science and Engineering, REVA University, Bengaluru, India

2 Department of Computer Science and Engineering, College of Engineering and Technology, Faculty of Engineering and Technology, SRM Institute of Science and Technology, Vadapalani Campus, Chennai, India

3 Department of Computer Science and Engineering, Saveetha Engineering College, Chennai, India

4 Department of CSE , Chaitanya Bharathi Institute of Technology, Hyderabad, India

cosine similarity. The proposed scheme highlighted and addressed the critical role of the KL divergence-based cluster ensemble factor in improving clustering stability and robustness. For prediction, the enhanced sqrt-cosine similarity was used to calculate successful related neighbor users. The performance of Recommendation is improved when KLD-FCM is combined with improved sqrt-cosine similarity.The proposed scheme's empirical work on the Movielens dataset in terms of MAE, RMSE, SD, and Recall were found to be superior in recommendation accuracy compared to traditional approaches and some non-clustering based methods recommended for study. With the specified number of clusters, it is capable of providing accurate and customized movie recommendation systems.

Keywords Recommendation systems · Content-based filtering · Knowledge base · Fuzzy C-means algorithm

1 Introduction

A Recommendation System (RS) is a filtering program that enables consumers to evaluate product recommendations for online purchases and provide information into products that they are interested in. In recent years, the extensive advancement of science and technology has resulted in vast amounts of digital knowledge being accessible via the internet. As a result of the overabundance of data, users are unable to obtain accurate taste information. This problem of information overload can be solved with RS, which filters out irrelevant information and only recommends related things to users.RS's knowledge filtering system assists users in making decisions in difficult situations by scanning vast data for items of interest. It also offers a personalized proposal based on the user's desires and preferences. RS is widely used in e-learning, e-shopping, e-tourism, e-business, e-government, and social networking sites like Facebook. Articles from research, books, music, news, films, DVDs/CDs, and other e-shopping products recommended by RS for their clients [1, 2].

To create a working recommendation system, a large amount of data must be collected. RS accepts a variety of inputs, both explicit and implicit. The explicit ratings of 1 to 5 given by users for their preferences in the items they purchase make up the covert reviews. User actions such as accessing and navigating past websites, click and search logs provide implicit input. Demographic data is another addition to RS. This information index was developed for each client who visits the site. Following this point, the data is filtered to obtain sufficient data for customer/user suggestions.Content-based filtering, collaborative filtering, and hybrid filtering technology are examples of filtering algorithms that would be more suitable for the recommending engine [3, 4]. The collection of data and application of recommendation filtering methods yield a set of recommendations that comply with the procedures to be considered in a recommendation system's calculation. In general, two types of performance are predicted and suggested. The ranking items for which the target consumer has not been rated are forecasted by prediction. On the basis of the forecasted ratings, the recommendation recommends the top-n recommendations to the target consumer, where each item does not include an evaluation value for these top-n recommendations. The consumer should be ecstatic with the performance of the recommender [5].

A framework that offers good and helpful recommendations for its own users requires the use of appropriate and reliable recommendation techniques. The content-based technology employs a domain-based algorithm that focuses on analyzing the characteristics of predictive posts. When documents such as blogs, magazines, and news are recommended, it is the most effective material filtration technique. The user profile recommendation in the Content-Based Filtering (CBF) approach is based on characteristics extracted from the content of items checked by the user.Objects that are mostly associated with positive items are recommended to the customer. CBF employs a variety of models to identify similarities between documents in order to generate useful recommendations. To form the relationship between different documents within the Corpus, it could use a Vector Space Model or a probabilistic model like the Naive Bayes Classification method, Decision Trees, or Neural Networks. It's also possible to use the Vector Classification method. These approaches make recommendations based on the underlying model's statistical analysis or machine learning techniques [6].

Other users' profiles aren't needed for content-based filtering because they don't affect the recommendation. Furthermore, as the user profile changes, CBF's approach adjusts its recommendations in a very short time. The key drawback of this method is that it necessitates thoroughly informing and explaining the characteristics of the objects in the profile. CF is a domain-agnostic content prediction technique that can't be easily or reliably classified by metadata like film or music. Filtering by working together creates a database of user preferences (user-item matrix). It then matches individuals with relevant interests and preferences to make recommendations.This group of people is creating a social network. A consumer receives recommendations for items that he hasn't yet rated but that other users in his field have given high marks to. CF may make recommendations in the form of forecasts or recommendations. Collaborative filtering has a range of benefits over CBF, including the ability to be used in fields where object content is scarce and computer system content is difficult to assess (such as opinions and ideals). CF technology should provide persuasion recommendations so that things that are useful to the user can be recommended, even though the user profile does not include the content [9].

Hybrid filtration strategies combine various recommendation approaches to boost device optimization and prevent some of the drawbacks and issues that come with pure recommending systems. Since one algorithm can overcome its drawbacks with another algorithm, a combination of algorithms is built to make recommendations more accurately and efficiently than a single algorithm. Using various recommendation models, a combined model will eradicate the flaws of a single process.Separate algorithms can be applied and the results merged, content-based co-filters can be used, content-based collaborative filters can be used, and a single recommendation system can be developed to bring all methods together [10].

The most promising products from which consumers can choose are included in the recommendation question. Some well-known approaches for solving the problem of scalability under model-based collaboration filters are clustering-based approaches. Predominantly, the majority of CF-based clustering strategies have relied on K-means and Fuzzy C-means clustering, which lack the ability to pick a relevant clustering core, lowering the predictive efficiency. As a result, there are trade-offs between scalability and predictive efficiency.Improved clustering techniques were suggested in the study to recognize agreed problems due to scalability. Many previous studies have shown that clustering-based CF systems (CF combined with clustering algorithms) are a promising schema for providing accurate personal recommendations and solving large-scale problems [8]. Fuzzy C-Means is a soft clustering approach that allows each individual data to be allocated to multiple clusters based on different membership degrees. They also concluded that good clustering-based CF performance is dependent on appropriate clustering techniques as well as the dataset's design. The analysis shows that the Clustering algorithm's stability and robustness need to be improved in order to achieve critical accuracy in the process of movie recommendation to target users. The following are the limitations of the current method:

- Fuzzy-C refers to a lack of ability to choose the initial cluster Center point, which can lead to a local optimum solution and affect clustering accuracy. In certain cases, the obtained clusters can be impractical, affecting the Recommendation outcomes.
- To get a good grouping of data, most current clustering algorithms require the configuration of some parameters.
- The disadvantage of FCM is that it has a higher error rate and needs further iterations to obtain well-framed clusters.
- Because of the prejudices and assumptions that each clustering algorithm contains, applying a single clustering method generally results in inconsistent results.

The Fuzzy C Means clustering with KL divergence is suggested to solve the aforementioned limitations. The research work's contributions have the following features.

- To prevent the drawbacks of a bad initialization, Ensemble FCM clustering is used to divide users into separate groups.
- The Ensemble Fuzzy C-Means clustering methods use a KL divergence-based cluster ensemble factor to improve the stability and accuracy of the clustering process, resulting in successful clustering with the goal of focusing on better performance results during movie recommendation.
- A better approach is to treat the membership vector as a discrete probability function, with the statistical distance, such as KL divergence, serving as the similarity metric.
- For active users, the enhanced sqrt-cosine similarity is often used to find the most powerful nearest neighbors.
- Reduce the problem of scalability.

The latest analysis of RS methods is summarized in Section 2. In addition, work on various RS is discussed in this chapter. The proposed framework model for the hybrid video recommender is defined in Section 3. The quality of predictions as measured by the assessment metrics is also stated. Section 4 brings the analysis to a close by highlighting the algorithms that aided in the achievement of the objectives and promoted the desired outcome. The study's limitations have been established, and potential research directions have been summarized.

2 Related study

Recommend systems use data mining and predictive algorithms to predict user preferences among the vast array of images, goods, and services available. The rapid growth of knowledge on the Internet, as well as the number of visits to websites, are posing significant challenges to system recommendations. The development of precise recommendations, the successful management of a large number of recommendations, and the large number of system members are all examples of these challenges.As a result, new system recommendation technologies are needed that can produce high-quality recommendations quickly, even for large data sets. There are numerous methods and algorithms for data filtering and recommendation. This section provides a brief overview of recent system-related studies in the literature.

Collaborative filtering is a widely used and relevant technology that makes predictions and suggestions based on the ratings and actions of other system users. The key premise of this strategy is that the user can pick and consolidate views from other users in order to understand his choice for the active user. Memory-based CF algorithms generate a forecast for the entire database or a subset of the user-item database. Each consumer is a member of a group of people who share similar interests. A prediction of a new user's tastes for new objects can be rendered by identifying related neighbors (or active users).Memory-based collective filtering has a number of drawbacks, the most significant of which is that it must use the entire database every time it predicts something, making it extremely sluggish in memory. If the rating matrix is so broad that many people use it, the issue becomes serious. Computing resources are depleted, and device performance suffers as a result, making it impossible for the system to respond quickly to user requests [11].

The model-based approach learns a model to boost collaborative filtering technology efficiency using prior scores. The model could be built using machine learning or data mining techniques. Since these approaches rely on pre-computed models, they can quickly suggest a large number of items and have been shown to yield results that are comparable to neighborhood-based recommendation techniques. Dimension reduction, for example, includes techniques including singular decomposition, matrix completion, latent semantic approaches, regression, and clustering.Content-based filters recommend items on a user's item profile and user profile. When an account is created and the framework is first used, these types of profiles are created. As a result of the user's interaction with the system, a better user profile is developed. If a user likes an object in the past, the CBF scheme assumes that the user would like similar things in the future [7]. The most powerful filtering technique is used in information documents such as web sites, journals, and news. To produce meaningful recommendations, CBF employs a variety of models to detect correlations between documents. Model-based vector space models can be used to represent the relationships between various documents within a corpus, such as reverse frequency or probabilistic models like the Naive Bayes Classification, Decision Trees, or Networks. Object metadata is used in the filtration mechanisms. Before users can get a recommendation, they'll need a large collection of items and a well-organized user profile. As a result, the effectiveness of CBF is dependent on the availability of descriptive data. Over-specialization is another major problem with the CBF methodology. Users can only get suggestions that are close to their own [12].

In certain applications, hybrids of various types outperformed individual algorithms. When the algorithms in question cover a wide range of use cases or aspects of the data set, hybrids can be particularly useful. The suggestion has been suggested to be implemented using a range of approaches, including material-, collaborative-, knowledge-based, and other techniques.Every form of recommendation has its own set of strengths and weaknesses. In order to improve efficiency, these strategies were often combined into hybrid recommenders. The hybrid recommender method has a higher level of complexity and implementation costs [14].

The Knowledge Base (KB) suggestion suggests things based on user experience, artifacts, and/or user relationships. In most cases, KB recommendations maintain a knowledge base describing how a particular item serves the needs of a specific person, which can be carried out based on inferences about the relationship between a user's need and a potential recommendation. The semantic similarity between objects can be calculated using the domain ontology. Social recommendation services are an integral part of everyday life on social media.Every minute, users on social media exchange details. Social advisor programs are designed to help people understand what they really want by reducing the amount of information available on social media. They want to help people on social networking sites like Facebook, Twitter, YouTube, Flickr, and Weibo by providing them with tweets and profiles that meet their needs [13].

Systems that are recommended are those that can analyze past user habits and make suggestions for current issues. Simply put, information on similar behavior, remarks, and users will be used in the RSs to try to define the user's thought style in order to assess and suggest the user's taste as the most appropriate and near object. Many of the methods in the RSs are used to make recommendations that are as accurate as the users need.As a result, several models for RSs clustering existing data have been presented in order to efficiently process data with large volumes. Given the goals, special characteristics, and relationships between data in the RSs, using an effective data cluster approach to efficiently process data in subsequent steps and produce more reliable suggestions has always been seen as an important research area for developing these systems [15].

Collaborative filtering recommendation systems are extremely useful for a wide range of online activities, including e-commerce. However, there are significant issues, especially in scalable and dynamic scenario implementations where new users, objects, and ratings are added frequently. Scalability refers to a system's, network's, or process's ability to control or expand its capacity to accommodate development. For example, a device is called escalable if it can increase its total power under increased load as resources (typically hardware) are added.Dimension reduction-based approaches address scalability problems such as SVD, MF, clustering, and so on. In short, cluster-based approaches retain the advantages of low computation cost (for searching candidates) over memory-based approaches as models for dimension reduction [16, 17].

To improve the performance of the recommendation, related users are grouped together based on their interests. Clustering was used to quickly locate a user's neighbors. By reducing the size of the original data to more manageable partitions, clustering systems can react quickly. Clustering, in particular, boosts the scalability and accuracy of recommendation systems. Despite the advantages of low machine costs (for searching candidates) over memory-based and SVD methods, MF methods remain cluster-based methods.One of the most serious problems with a recommending method is scarring, and data sparing has an effect on the accuracy of the recommendation. In general, machine data like Movie Lentils is interpreted as a user-item matrix made up of films, which increases matrix dimensions and sparsity since the user and items are no longer used. Most users do not rate most items, and there are few available ratings. The key explanation for this is a lack of knowledge. In order to condense the user's object matrix, the reduction of dimension addresses the issue of scarcity by excluding non-representative or insignificant users or objects. However, during the reduction process, potentially valuable information is lost.However, some potentially valuable information can be lost during the reduction process [18, 19].

Collaborative filters create this problem because they depend on the rating matrix in most cases. Many researchers have attempted to address this problem, but more research is still required in this area. Most recommendation systems on the major electronic commerce platforms have been influenced by the long tail effect in some way. Since accuracy-focused recommender systems tend to recommend common goods, recommending items with few ratings is critical (long tail items). Popular products that are likely to be less helpful to users can be easily recommended using detailed recommendation algorithms.To assess the ability of systems to recommend unpopular goods, the assessment metrics diversity and innovation have been added. Recommending long tail artifacts can result in the recommendation's precise results being lost. As a result, a recommendation process must be developed that recommends controversial products while minimizing accuracy loss. Several guidelines have recently been proposed to strike a balance between precision, diversity, and novelty [20–22].

Existing CF clustering method algorithms are ineffective at improving RS efficiency and addressing scalability issues. The recommended performance has an effect on the efficiency of the clustering procedures [23, 24]. As a result, there is a lack of precision and coverage, which makes clustering-based approaches in recommender systems difficult to use in practice. To improve the recommendation's performance, better methods for optimizing the above problem are needed.

3 Proposed system model

3.1 KL divergence based ensemble fuzzy C means clustering

Cluster-based CF has been shown to address scalability issues while also improving the consistency of recommendation outcomes in recent years. The aim of clustering algorithms is to group objects into clusters with the shortest distance between them in order to find objects that are identical. Clustering strategies will typically group a large number of users into various clusters based on their rating similarity in order to locate "like-minded" neighbors.One of the most widely used clustering methods is fuzzy clustering. To get a decent grouping of data, most current FCM algorithms require the specification of some parameters. As a result, the Fuzzy cluster ensemble solution usually prevents the drawbacks of a bad initialization.

3.1.1 Fuzzy C means clustering algorithm

1. Initialize Membership matrix M with random data points.
2. Fuzzy cluster center is calculated C
3. Calculate the objective function F = M*($Xi-Cj$)
4. For every iteration fuzzy Membership is updated by using M=Σ CK Where k is the number of clusters
5. The iteration will stop when (k+1) − (k) < termination criterion

Ensemble clustering blends a dataset's various simple partitions into a more stable and robust one. The basic concept behind the Ensemble Fuzzy C Means cluster method is to apply the clustering method to the data several times (rather than only once) and then merge the results into a single partition. Ensemble clustering takes a collection of data partitions as input. The cluster ensembles are divided into two parts. The ensemble clustering generator is one, and the consensus function is another.The first section focuses on generating more diverse clustering results, while the second section focuses on seeking a good consensus feature to increase the results' accuracy. Homogeneous ensemble FCM clustering is used in the first component of cluster ensembles. The term "homogeneous ensemble" refers to the use of multiple runs of a single clustering algorithm (fuzzy c-means algorithm) with different initializations and fuzzy parameter values. Several soft partitions of the data are obtained at the end of the first stage of the ensemble method as a result of several runs of the algorithm(s). The aim of this is to improve the accuracy and consistency of fuzzy cluster analysis procedures. Soft ensembles are characterized by the concatenation of membership probability distributions in the second part of cluster ensembles. The obtained partition will be combined in the second stage using a KL divergence-based objective function to produce a single final partition. The Kullback–Leibler (KL) divergence was then used to describe a distance measure between two instances. The similarity of a membership vector to a cluster center is measured by FCM using squared Euclidean distance.This is ineffective in situations where a data's membership in all clusters normally equals one. A better approach is to treat the membership vector as a discrete probability function, with the statistical distance, such as KL divergence, serving as the similarity metric. This algorithm is identical to fuzzy c-means, with the exception that it employs the KL divergence to treat memberships as discrete probabilities.

The data from the User Item is first categorized using homogeneous fuzzy clustering methods. After that, a fuzzy KL divergence-based objective function aggregates the soft clustering effects.

3.2 Improved Sqrt-cosine similarity

Improved Sqrt-Cosine Similarity (ISC) is a modern similarity measure that uses Hellinger distance and is based on sqrt cosine similarity. Hellinger distance (L1 norm) is a much better metric for high-dimensional data mining applications than Euclidean distance (L2 norm). In terms of implementation, the ISC is very similar to cosine similarity, and it outperforms other similarity measures in high-dimensional results.The enhanced sqrt-cosine similarity determines how close two users are. For High Dimensional data, the Hellinger distance-based Similarity is more accurate. The KLD-FCM with enhanced sqrt-cosine similarity outperforms current systems in terms of recommendation performance.

3.3 Proposed KLD-FCM based movie recommendation scheme

For improving movie recommendation methods (KLD FCM-RS), a kullback–leibler divergence-based fuzzy c-means clustering is proposed, and the steps involved in KLD FCM-RS are discussed in detail. The aim of the proposed method is to develop a Collaborative Movie Recommender framework that can solve scalability problems while also improving prediction accuracy in terms of MAE, Precision, Recall, and Speed. The proposed KLD-FCM based Movie Recommendation Scheme is architecturally depicted in Fig. 1 as follows.

It is divided into two phases: offline and online. The User Item Rating Matrix derived from the used Movie Lens Dataset is used as a possible input during the offline process. Then, over the extracted user Item scores, a method of different homogeneous Fuzzy C means clustering is applied to divide the users into different classes. Furthermore, KL Divergence-based cluster ensemble FCM is used to combine the various FCM clustering findings in order to generate efficient single User clusters. The nearest cluster estimation for Active consumer is computed using the Euclidean distance method in the online process.The active user's nearest neighbors in his or her nearest cluster are then found using an enhanced sqrt-cosine similarity tool. Finally, the top list of recommended movie items is calculated in an online mode by determining the movie items that are most frequently recommended by the context's neighborhood users. Leibler–Kullback For the Recommender method, divergence-based Fuzzy C-Means clustering with enhanced sqrt-cosine similarity worked well. Three possible steps are included in this proposed KLD-FCM:

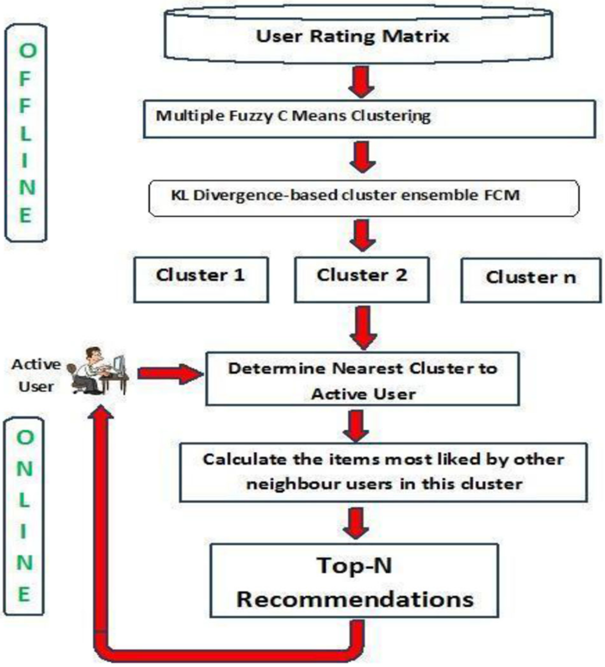


Fig. 1 Architectural view of the proposed KLD-FCM Scheme

3.3.1 Procedure

Step 1 (FCM-KLD) FCM-KLD is divided into two phases. Multiple Fuzzy C Means Clustering is the first step. The user data ratings from the Movie Lens data set are used as input in this phase of the clustering process. Over the User Item scores, apply three different homogeneous Fuzzy C Means clustering methods with different initializations. By executing the FCM several times for each initialization with different fuzzy parameter values, homogeneous FCM clustering is used to create several partitions with different random initializations (here 1.5, 2 and 2.5 are used).**Input:** User Rating Matrix **Output:** Three Clustering results.

Table 1 shows a snapshot of the User object rating matrix from the Movie Lens dataset for 5 users on 5 movies, where U1-U5 are users and M1-M5 are objects (movies). The value 1–5 represents the user's likelihood rating for a specific film. The value '0' denotes that the consumer has not rated (or seen) the film. The recommender framework identifies unrated values and suggests the top N films to the consumer Tables 2, 3, 4, and 5.

Step 2 (determine the nearest cluster to active user) After clustering the users into various clusters, the Euclidean distance approach is used to determine the nearest cluster to Active User.

$$sim_i(Cent_i, U) = \sum_{j=1}^{d} \left(Cent_{i,j} - U_j\right)^2 \tag{1}$$

$Cent_i$ is the centroid of 'i' th cluster, U is the Active User Profile.
d is the dimension of data(Number of Attribute).
$Cent_{i,j}$ is the jth attribute of centroid profile in cluster i.

Table 1 Example of rating matrix from Movie Lens dataset

	M1	M2	M3	M4	……	M_m
U1	5	3	4	3	……	3
U2	4	?	?	?	……	1
U3	?	?	?	?	……	0
U4	5	?	?	2	……	?
….	……	……	……	……	……	……
U_n	4	3	?	?	……	1

Table 2 MAE comparison analysis

Cluster size	5	10	15	20	25
Proposed	1.2	1.2	1.15	1.15	1.14
FCM	1.3	1.29	1.28	1.28	1.27
FCMBAT	1.4	1.39	1.38	1.38	1.37
COA	1.5	1.49	1.48	1.48	1.49

Step 3 (using improved sqrt-cosine similarity, determine the top N recommended movies) Improved sqrt-cosine similarity is used to calculate Active User's nearest neighbors. The following formula is used to determine how close users u1 and u2 are.

$$sim(u1,u2) = \frac{\sum_{i=1}^{m}\sqrt{R_{u1,i}R_{u2,i}}}{\sqrt{\left(\sum_{i=1}^{m}R_{u1,i}\right)}\sqrt{\left(\sum_{i=1}^{m}R_{u2,i}\right)}} \tag{2}$$

m Set of common items rated by user u1 and user u2.
$R_{u1,\,i}$ is the rating given to item 'i' by user u1.
$R_{u2,\,i}$ is the rating given to item 'i' by user u2.

Table 3 RMSE comparison analysis

Cluster Size	5	10	15	20	25
Proposed	0.78	0.77	0.76	0.76	0.75
FCM	0.81	0.8	0.79	0.78	0.77
FCMBAT	0.84	0.84	0.84	0.84	0.81
COA	0.87	0.86	0.86	0.85	0.83

Table 4 Recall analysis

Cluster size	5	10	15	20	25
Proposed	25.28	25.68	27.24	27.8	28.56
FCM	23.12	24.68	25.45	25.12	25.89
FCMBAT	20.12	22.68	23.45	23.12	24.89
COA	18.9	19.89	20.35	20.02	21.67

Table 5 Accuracy analysis

Cluster size	5	10	15	20	25
Proposed	89.3	88.2	88.2	87.1	87
FCM	86.3	86.6	86.3	85.4	85
FCMBAT	84.8	84	83.7	83.5	84.3
COA	82.5	82.4	82.3	82.1	81.67

The Hellinger distance is used to compute the similarity between two vectors in the enhanced sqrt-cosine similarity. This phase is essential for comparing each individual user's rating to the ratings of other users in the clusters. Finally, the top list of recommended movie items that could be suggested to an active user at any time is calculated based on the movie items that are most often recommended by the context's neighborhood users. The movies are recommended to target users who are most likely used by other neighbor users and are not used by him/her during the Recommendation process.The weighted average of the ratings of items in the same cluster neighbor's is used to predict the ranking of unrated items for active users, and then the top-N suggestions list is sent to the active user. The rating of unrated movie (item) 'i' for an active user 'a' is predicted by $P_a(i)$

$$Pa(i) = \underline{R}_a + \frac{\sum_{N\varepsilon Cx} Sim(a,N) \times \left(R_N\ (i) - \underline{R}_N\right.}{\sum_{N\varepsilon Cx}(|Sim\ (a,N)|)} \tag{3}$$

Where,

a	Active User.
$\underline{R}_a$	Average of active user a.
Cx	set of nearest neighbors of active user a belonging to one common cluster.
$\underline{R}_N$	Average rating score given by active user's neighbor N.
$Sim(a,\ N)$	similarity between active user a and Neighbor.

4 Experimental design

Experiments using the publicly accessible Movielens dataset are used to equate the performance of the proposed KLD-FCM with that of baseline recommendation system The Movielens data set used to equate the proposed KLD-FCM scheme to the compared COA, FCM-BAT, FCM, and ICF contains 10,00,000 ratings, with 850 users theoretically rating them. This Movielens data set contains reviews ranging from approximately 1000 to 1513 movies, each scored on a scale of 1 to 5. By partitioning the entire Movielens data set using the k-cross validation process, the performance of the proposed KLD-FCM approach is investigated. The findings were evaluated using 5-fold cross validation. The original dataset is divided into five equal subsets. One is used as a test set (20%), while the other is used as a training set (80%). The procedure is repeated five times, with a different test set selected each time, and the average results recorded.In terms of MAE and RMSE, the proposed method was compared to non-clustering methods such as Basic CF (BCF), User-Based CF (UBCF), SVDM (a variant of Single Value Decomposition (SVD) that uses batch learning with a learning momentum), and RSVD (Regularized SVD model). Various collaborative approaches are selected from the literature to verify the proposed method's role in comparison to other cluster-based techniques.

ICF(integrated Collaborative Framework) The merits of the item k-NN algorithm were used to propose an Integrated Collaborative Framework (ICF). This ICF also provided classification restrictions, ensuring that only potential rules are used during the collaborative filter-based categorization of user ratings.

UPCC (user based CF Pearson correlation coefficient based CF) For the Collaborative filtering recommender scheme, the Pearson Correlation Coefficient test is used to determine how closely two users are related.

FCM (fuzzy C means) The performance of User-Based Collaborative Filtering with Fuzzy C Means is compared to that of other clustering methods such as K-means and Self-Organizing Maps (SOM).

FCMBAT The Fuzzy C-Means and BAT-based Movie Recommendation Scheme (FCM-BAT) is an integrated Fuzzy C-Means and BAT-based Movie Recommendation Scheme for promoting efficient and collaborative recommendation to the target users. This FCM-BAT was proposed to address scalability issues and improve the clustering process, with the aim of improving the consistency of the recommendation process.

COA(cuckoo Optimization Algorithms) Furthermore, a possible movie recommendation system based on k-means and COA is proposed in order to improve the rate of recommendation accuracy when using the Movielens dataset.

4.1 Mean absolute error

The suggested method used to measure MAE can be seen in Fig. 2 for different numbers of neighbors.

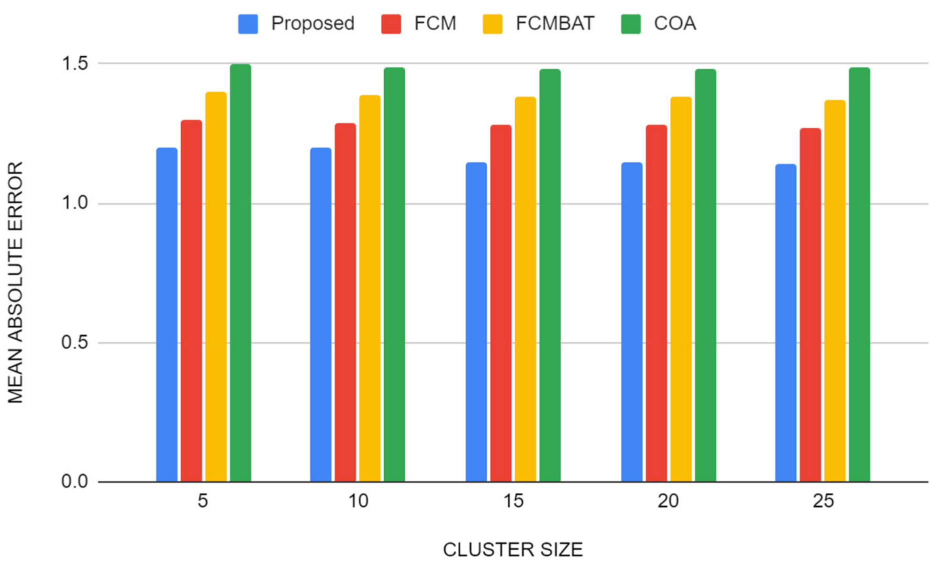


Fig. 2 MAE comparison analysis

Figures 2 and 3 show the contrast of the proposed FCM KLD with current Collaborative Recommender framework methods in terms of MAE and RMSE for different numbers of clusters. The proposed scheme's MAE and RMSE are proving to be significantly lower than those of current schemes. As a result, the proposed scheme's MAE is lower than the COA, FCM-BAT, FCM, and ICF approaches. The proposed scheme's RMSE is also tested to be substantially lower than the baseline schemes under consideration.

Figure 4 graphically depicts the contrast of the proposed KLD-FCM with current Collaborative Recommender framework approaches in terms of recall. It highlights the efficiency of the proposed KLD-FCM scheme as measured by Recall for a variety of cluster sizes. The proposed KLD-FCM scheme's Recall value is determined to be excellent as compared to baseline methods, since the KL divergence factor's guidance in the clustering phase is responsible for the majority of progress.

Figure 5 depicts a graphic comparison of the proposed KLD-FCM with current Collaborative Recommender framework approaches in terms of accuracy. The performance of the proposed method is significantly better than that of current methods. Since it takes advantage of the advantages of KL Divergence Fuzzy C, Clustering with enhanced sqrt-cosine similarity is possible.

5 Conclusion

The KL Divergence Fuzzy C means Clustering with improved sqrt-cosine similarity Recommender framework (KLD-FCM) is proposed to solve the CF technique's scalability problem. For successful clustering, Kullback–Leibler Divergence-based Fuzzy C-Means clustering is suggested, with the aim of focusing on greater accuracy during movie recommendation. The proposed KLD-FCM scheme is described as a trustworthy contribution that significantly improves the ability of movie recommendation by virtue of the KL divergence dependent Fuzzy C-Means clustering mechanism and enhanced sqrt-cosine similarity.The proposed

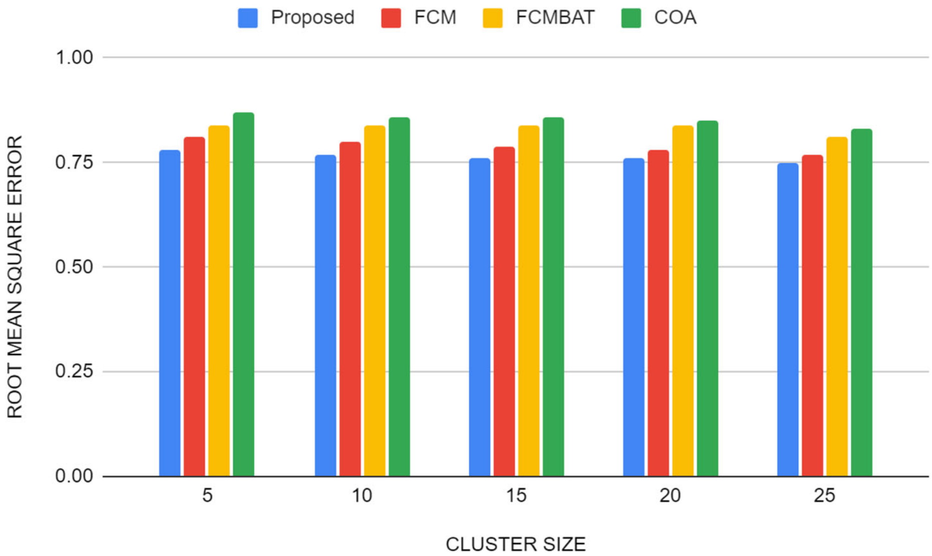


Fig. 3 RMSE comparison analysis

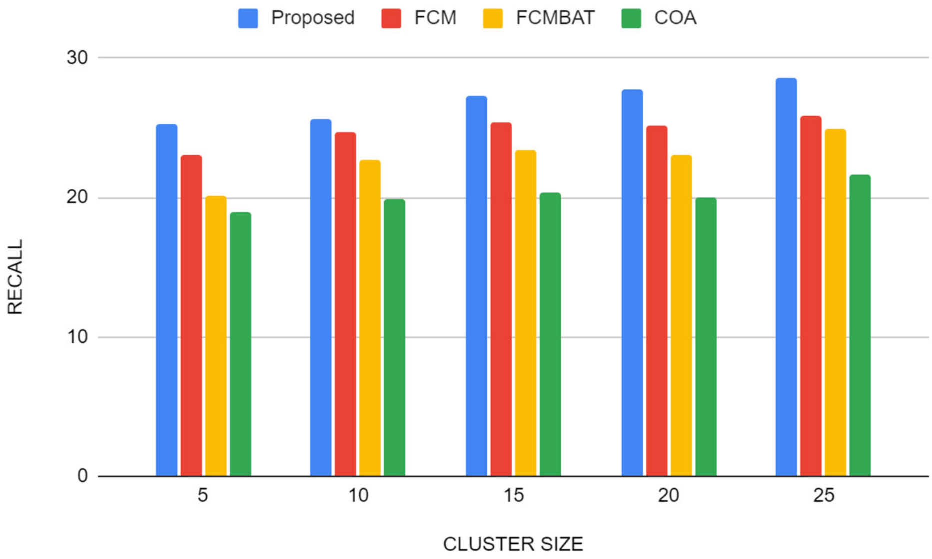


Fig. 4 Recall comparison analysis

scheme emphasized and presented the critical role of the KL divergence-based cluster ensemble factor in improving clustering stability and robustness. For prediction, the enhanced sqrt-cosine similarity was used to calculate successful related neighbor users. The performance of Recommendation is improved when KLD-FCM is combined with improved sqrt-cosine similarity.The proposed KLD-FCM scheme was found to be superior in recommendation Accuracy compared to the COA, FCM-BAT, FCM and ICF approaches, as well as some non-clustering based methods considered for study, when tested on the Movielens dataset in terms of MAE, RMSE, SD, and Recall. With the specified number of clusters, it is capable of

Fig. 5 Accuracy analysis

providing accurate and customized movie recommendation systems. In future work, the proposed design has to be tested with different datasets.

Declarations

Conflict of interest The author(s) propose a clear no conflict of interest involved in this research work in form of publication in this Journal Multimedia Tools and Applications.

References

1. Antony Vijay J, Anwar Basha H, Arun Nehru J (2021) A dynamic approach for detecting the fake news using random forest classifier and NLP. In: In computational methods and data engineering. Springer, Singapore, pp 331–341
2. Asadi E & Charkari N (2012). Video summarization using fuzzy c-means clustering. ICEE 2012 - 20th Iranian Conference on Electrical Engineering. 690–694. https://doi.org/10.1109/IranianCEE.2012.6292442
3. Basha SM, Rajput DS (2019). Survey on evaluating the performance of machine learning algorithms: past contributions and future roadmap. In deep learning and parallel computing environment for bioengineering systems. Academic Press, Cambridge, pp 153–164
4. Clement J. (2020). Impact of recommendation engine on video-sharing platform -YouTube. https://doi.org/10.13140/RG.2.2.15746.50882
5. Cui L & Dong L & Fu X & Wen Z & Lu N & Zhang G. (2016). A video recommendation algorithm based on the combination of video content and social network: CONTENT AND SOCIAL NETWORK BASED VIDEO RECOMMENDATION. Concurrency and Computation: Practice and Experience. 29: https://doi.org/10.1002/cpe.3900.
6. Davidson J, Liebald B, Liu J, Nandy P, Vleet T, Gargi U, Gupta S, He Y, Lambert M, Livingston B, Sampath D (2010). The YouTube video recommendation system. 293–296. https://doi.org/10.1145/1864708.1864770
7. De Vriendt J, Degrande N, Verhoeyen M (2011) Video Content Recommendation: An Overview and Discussion on Technologies and Business Models. Bell Labs Tech J 16:235–250. https://doi.org/10.1002/bltj.20513
8. Deldjoo Y. (2019). Enhancing video recommendation using multimedia content. https://doi.org/10.1007/978-3-030-32094-2_6
9. Deldjoo Y, Elahi M, Quadrana M, Cremonesi P (2015). Toward Building a Content-Based Video Recommendation System Based on Low-Level Features. https://doi.org/10.1007/978-3-319-27729-5
10. Deldjoo Y, Elahi M, Cremonesi P, Garzotto F, Piazzolla P, Quadrana M (2016) Content-Based Video Recommendation System Based on Stylistic Visual Features. J Data Semant 5:1–15. https://doi.org/10.1007/s13740-016-0060-9
11. Deldjoo Y, Schedl M, Cremonesi P, Pasi G (2020) Recommender Systems Leveraging Multimedia Content. Comput Surv 53:1–38. https://doi.org/10.1145/3407190
12. Gupta M, Thakkar A, Gupta V, Rathore DP (2021). Movie Recommender System Using Collaborative Filtering. 978–979
13. Homann L, Martins D, Vossen G, Kraume K (2018) Enhancing traditional recommender systems via social communities. Vietnam J Comput Sci 6. https://doi.org/10.1142/S2196888819500040
14. Huang Y, Cui B, Jiang J, Hong K, Zhang W, Xie Y (2016). Real-time Video Recommendation Exploration. 35–46. https://doi.org/10.1145/2882903.2903743.
15. Kamran M, Shah SS, Baig MN, Khan RH (2020). A movie recommender system by combining both content based and collaborative filtering algorithms
16. Khadse VP, Basha SM, Iyengar N, Caytiles R (2018) Recommendation engine for predicting best rated movies. Int J Adv Sci Technol 110:65–76
17. Lu W & Chung FL (2016). Computational Creativity Based Video Recommendation. 793–796. https://doi.org/10.1145/2911451.2914707.
18. Mercanoglu O & Yıldırım Z (2017). Video Recommendation System Using Collaborative Filtering
19. Mohamed A, Sherif A, Osama F, Roshdy Y, Hassan MA, El Ashmawi WH (2020). A new challenge on video recommendation by content. https://doi.org/10.1109/ICCES48960.2019.9068169.
20. Patil, Lalit. (2016). Fuzzy C means clustering MATLAB code. https://doi.org/10.13140/RG.2.1.3924.9046.

21. Ramezani M, Yaghmaee F (2016) A novel video recommendation system based on efficient retrieval of human actions. Physica A: Statistical Mechanics and its Applications 457. https://doi.org/10.1016/j.physa.2016.03.101
22. Shah P, Sanghvi S (2020) Video Recommender System
23. Tohidi N, Dadkhah C (2020) Improving the performance of video Collaborative Filtering Recommender Systems using Optimization Algorithm. Int J Nonlinear Anal Appl (IJNAA) 11:283–295. https://doi.org/10.22075/IJNAA.2020.19127.2058
24. Zhou X, Chen L, Zhang Y, Cao L, Huang G, Wang C (2015). Online Video Recommendation in Sharing Community. 1645–1656. https://doi.org/10.1145/2723372.2749444.

Expert Systems / Volume 40, Issue 5 / e13235

ORIGINAL ARTICLE

Optimizing CNN-LSTM hybrid classifier using HCA for biomedical image classification

Ashwini Kumar Pradhan, Kaberi Das, Debahuti Mishra, Premkumar Chithaluru

First published: 25 January 2023
https://doi.org/10.1111/exsy.13235

Abstract

In medical science, imaging is the most effective diagnostic and therapeutic tool. Almost all modalities have transitioned to direct digital capture devices, which have emerged as a major future healthcare option. Three diseases such as Alzheimer's (AD), Haemorrhage (HD), and COVID-19 have been used in this manuscript for binary classification purposes. Three datasets (AD, HD, and COVID-19) were used in this research out of which the first two, that is, AD and HD belong to brain Magnetic Resonance Imaging (MRI) and the last one, that is, COVID-19 belongs to Chest X-Ray (CXR) All of the diseases listed above cannot be eliminated, but they can be slowed down with early detection and effective medical treatment. This paper proposes an intelligent method for classifying brain (MRI) and CXR images into normal and abnormal classes for the early detection of AD, HD, and COVID-19 based on an ensemble deep neural network (DNN). In the proposed method, the convolutional neural network (CNN) is used for automatic feature extraction from images and long-short term memory (LSTM) is used for final classification. Moreover, the Hill-Climbing Algorithm (HCA) is implemented for finding the best possible value for hyper parameters of CNN and LSTM, such as the filter size of CNN and the number of units of LSTM while fixing the other parameters. The data-set is pre-processed (resized, cropped, and noise removed) before feeding the train images to the proposed models for accurate and fast learning. Forty-five MR images of AD, Sixty MR images of HD, and 600 CXR images of COVID-19 were used for testing the proposed model 'CNN-LSTM-HCA'. The performance of the proposed model is evaluated using six types of statistical assessment metrics such as; Accuracy, Sensitivity, Specificity, F-measure, ROC, and AUC. The proposed model compared with the other three types of hybrid models such as CNN-LSTM-PSO, CNN-LSTM-Jaya, and CNN-LSTM-GWO and also with state-of-art techniques. The overall accuracy of the proposed model received was 98.87%, 85.75%, and 99.1% for COVID-19, Haemorrhage, and Alzheimer's data sets, respectively.

CONFLICT OF INTEREST

The authors state that they have no known competing financial interests or personal ties that could have influenced the research presented in this study.

Open Research

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Kaggle at https://www.kaggle.com/search?q=covid+19+dataset .

REFERENCES

Abraham, M. K., & Chang, W.-T. W. (2016). Subarachnoid hemorrhage. *Emergency Medicine Clinics*, **34**(4), 901–916.

PubMed | Web of Science® | Google Scholar

Addeh, A., Khormali, A., & Golilarz, N. A. (2018). Control chart pattern recognition using rbf neural network with new training algorithm and practical features. *ISA Transactions*, **79**, 202–216.

PubMed | Web of Science® | Google Scholar

Addeh, J., & Ebrahimzadeh, A. (2012). Breast cancer recognition using a novel hybrid intelligent method. *Journal of Medical Signals and Sensors*, **2**(2), 95–102.

PubMed | Google Scholar

Aditi, M. K., & Poovammal, E. (2019). Image classification using a hybrid lstm-cnn deep neural network. *The International Journal of Engineering and Advanced Technology (IJEAT)*, **8**(6), 1342–1348.

Computational-Intelligence-Inspired Adaptive Opportunistic Clustering Approach for Industrial IoT Networks

Premkumar Chithaluru, Fadi Al-Turjman, Manoj Kumar, *Member, IEEE*, and Thompson Stephan, *Member, IEEE*

***Abstract*—The major issues and challenges of the Industrial Internet of Things (IIoT) include network resource management, self-organization; routing, mobility, scalability, security, and data aggregation. Resource management in IIoT is a challenging issue, starting from the deployment and design of sensor nodes, networking at cross-layer, networking software development, application types, environmental conditions, monitoring user decisions, querying process, etc. In this article, computational intelligence (CI) and its computing, such as neural networks and fuzzy logic, are used to tackle the challenges of resource management in the IIoT. The incorporation of the neuro-fuzzy technique into the IIoT contributes to the self-managing intelligence systems' self-organizing and self-sustaining capabilities, offering real-time computations and services in a pervasive networking environment. Most of the problems in IIoT are real-time based; they require fast computation, real-time optimal solutions, and the need to be adaptive to the situation of the events and data traffic to achieve the desired goals. Hence, neural networks and fuzzy sets would form appropriate candidates for implementing most of the computations involved in the issues of resource management in IIoT networks. A real-time testbed network is simulated and implemented on the Crossbow mote (sensor node) using TinyOS.**

***Index Terms*—Computational intelligence (CI), industrial IoT, mobile node, neuro-fuzzy technique, resource management, self-managing, self-organizing, self-sustaining.**

I. Introduction

THE INDUSTRIAL Internet of Things (IIoT) is dynamic in nature. The input traffic and other environmental data are fraught with uncertainty [1]. They are prone to unanticipated overloads and outages. Fuzzy logic looks to be a potential solution for addressing such critical features of IIoT [2], [3]. It provides a strong mathematical foundation for dealing with real-world imprecision and nonstatistical uncertainty. A thorough search and study of the literature demonstrates that current research on fuzzy logic in communication networks ranges from queuing, buffer management, and load management to routing, bandwidth allocation, network administration, and quantitative network performance evaluation [4].

IIoT is becoming more prevalent in the above-mentioned real-time operations and emerging applications. As a result, there is a need to investigate and comprehend system behavior in a variety of environments based on potential application requirements and issues, such as resource management, self-organization, routing, mobility, and scalability. Along with these issues and challenges in IIoT, the techniques should tackle the requirements of applications like real-time event detection, location-based monitoring, operating environment decisions, etc. [6], [7].

1) *Resource Management:* IIoT is subjected to a unique set of resource constraints. Some of them include limited hardware, networking, and support for software development. IIoT mobile nodes are unreliable, which causes network failures and faults more often. Similarly, there are no universally accepted standard routing protocols or network services. In such a resource-constrained environment, resource allocation in a distributed environment, their activation, data storage, computation, and preprocessing allocation are to be monitored [8]. The resource reservation for future applications and different events of occurrence based on the environmental parameters are to be reserved and allocated. Application priority, real-time requirements, and data-intensive task-based applications are to be synchronously and asynchronously monitored. Link bandwidth utilization, throughput, and bit error rate are the essential link parameters in the IIoT that are to be allocated, reserved, and monitored for better network resource management using computational intelligence (CI) methods.
2) *Scalability, Mobility, and Dynamic Network Topology:* The number of sensor nodes (n_i) placed to examine an IIoT environment for many envisioned applications could be in the hundreds or thousands. Depending on the application, the number might be in the millions. Scalability in a network environment occurs when an application grows; the network should be flexible

Manuscript received 29 January 2022; revised 28 March 2022 and 24 May 2022; accepted 8 December 2022. Date of publication 6 January 2023; date of current version 25 April 2023. *(Corresponding author: Manoj Kumar.)*

Premkumar Chithaluru is with the Department of Computer Science and Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad 500075, India (e-mail: bharathkumar30@gmail.com).

Fadi Al-Turjman is with the Research Center for AI and IoT, Faculty of Engineering, University of Kyrenia, 99320 Kyrenia, Turkey, and also with the Artificial Intelligence Engineering Department, AI and Robotics Institute, Near East University, 1403 Nicosia, Turkey (e-mail: fadi.alturjman@neu.edu.tr).

Manoj Kumar is with the Faculty of Engineering and Information Sciences, University of Wollongong in Dubai, Dubai, UAE (e-mail: wss.manojkumar@gmail.com).

Thompson Stephan is with the Department of Computer Science and Engineering, M. S. Ramaiah University of Applied Sciences, Bengaluru 560054, India (e-mail: thompsoncse@gmail.com).

Digital Object Identifier 10.1109/JIOT.2022.3231605

enough to allow this expansion to occur anywhere and at any time without interfering with network performance. The addition of extra nodes, malfunctioning, device failure, shifting impediments, mobility, and interference all cause frequent changes in the network architecture [9]. Despite these dynamics, the IIoT should provide resilient operation by responding to the changing network environment.

3) *Network Self-Organization:* Because of the application's nature, most n_i's placed in the network with no infrastructure. IIoT networks are often installed in a forest randomly; moreover, they are inherently unattended in most situations; once deployed, IIoT networks require no human involvement. In such a case, the nodes must self-organize to determine their connection and dispersion because the manual setup is not possible [10]. As a result, the network must be able to reconfigure itself frequently. The autonomous network functioning is a critical design problem as well as network self-organization is a difficult problem in IIoT [11].
4) *Quality-of-Service (QoS) Guarantee:* QoS typically refers to the perceived quality of the application/user. It is a standard that QoS must be satisfied while transmitting a data packet from its origin to its destination. The QoS needs end-to-end change in traditional IIoT applications. As a result, IIoT should provide real-time services as well as fundamental data delivery mechanisms [12]. IIoT research areas focus on end-to-end QoS, dependability, and application-specific QoS.

To overcome the aforementioned limitations, we proposed an adaptive opportunistic clustering approach (AOCA) based on CI for IIoT Networks to increase mobility support and extend IoT devices' network lifetime. The purpose of the AOCA schema is to enhance self-managing intelligent systems' ability to self-organize and self-sustain by providing real-time computations and services in an ubiquitous networking environment. Most of the IIoT problems are real-time in nature, requiring quick computing, real-time optimal solutions, and the ability to react to the circumstances of events and data traffic to accomplish desired results. Hence, neural networks and fuzzy sets would be the best choices for IIoT networks to overcome the issues of resource management.

The major contributions of this research paper are as follows.

1) An enhanced Neuro-Fuzzy technique is proposed to improve self-managing intelligent systems' ability to self-organize and self-sustain, offering real-time computations and services in a pervasive networking environment.
2) Neural networks and fuzzy sets were incorporated for each cluster's dynamic clustering and resource management. The concept entails using CI to create cluster routing in an IIoT network.
3) The proposed technique is merged with hop-field neural networks (HNN) to compute the best transmission path in-between source and destination.
4) Validate the energy optimization parameters, such as the first node dies (FNDs), last node dies (LNDs), and a number of packets send to cluster head (CH) and sink on the proposed and peer competing routing protocols like and adaptive ranking fuzzy-based energy-efficient opportunistic routing (ARFOR) scheme [23] and multilayer threshold cluster-based energy efficient low power and lossy networks (MECEE-LLNs) [24] for simulation and testbed implementation using the TinyOS simulator (TOSSIM).

The remainder of this article is arranged as follows. Section II covers a study of the literature on various peer routing protocols. Section III describes the operation of the proposed AOCA, while Section IV presents the results and discussions. Finally, Section V concludes the work by discussing future potential research.

II. Literature Review

Due to the increasing popularity of IIoT, many approaches are being implemented to improve its functionality. Many routing algorithms have been proposed and implemented over the years to improve the clustering and routing operations in IIoT networks. This section describes the existing clustering protocols and the design issues associated with them in the IIoT environment. A few routing algorithms are presented in this article based on a review of the literature.

Traditional protocols, such as [13], [14], [15], and [16], offered a two-phase clustering (TPC) strategy for energy conservation and delay-adaptive data collection in IIoT. In phase I, the algorithm divides the network into clusters, each with its CH, providing a direct link between cluster members and CH. In phase II, each cluster member seeks a neighbor within the cluster who is closer than the CH to establish an energy-saving data relay link. Depending on the needs stated by the users or application, the sensors employ either the direct link or the data relay link to send detected data.

The research works [17], [18], [19], [20], [21], [22], [23], [24], proposed clustering scheme for IIoT, for periodical data gathering applications. The approach selects CHs with more residual energy through local radio communication while achieving well CH distribution. They have also introduced a method to balance the load among the CHs.

A. Problems Identified in Literature

To improve network longevity, mobility, and self-organization in IIoT, the following sections highlight the research gaps and issues that have been identified.

1) In a few clustering algorithms, resource-rich nodes are designated as CHs. The disadvantage of these strategies is that most IIoT are homogeneous and resource-constrained. As a result, the approach is ineffective in various situations. Furthermore, even if a resource-rich node in a heterogeneous network can be discovered and designated as a CH, being CH for an extended period will rapidly deplete the node's power and induce node death.
2) CHs are chosen depending on a variety of factors, such as available resources, location, number of neighbors, and so on. CH selection algorithms can increase network

efficiency by identifying nodes that are more suited to be CHs. Furthermore, certain approaches are meant to react to unanticipated conditions by dynamically reselecting or replacing CHs with more relevant nodes.

Despite these research gaps, there is still inefficiency in dealing with a resource-constrained environment, resource allocation, and clustering. As a result, the proposed protocol employs CI, fuzzy logic, and HNN to improve clustering and choose nodes within each cluster to decrease energy usage.

III. Proposed AOCA Protocol

Improving operational efficiency in a resource-constrained environment is a challenging issue in IIoT. In such network dynamics and unpredictable behavior of network parameters, such as wireless media, tiny computing nodes, power, and applications requirements need to be analyzed. In this section, we proposed a technique for clustering in the network based on energy. Cluster analysis is the process to explore cluster formation, network topology characteristics, and application requirements. The n_i's are clustered and send the data to their CHs. CHs collect the data and send it to the sink. The proposed method is implemented using TinyOS by establishing a suitable IIoT.

Algorithm 1 contains the pseudocode for the static node parent selection procedure.

A clustering and cluster analysis technique is proposed for studying the behavior and operation of IIoT in terms of node and network parameters. There are application characteristics that have an impact on the resources available in IIoT. As a result, there is a need to understand how IIoT acts in a resource-constrained context in terms of application requirements.

CI has been implemented in our proposed clustering technique. CI is a collection of nature-inspired computational techniques and approaches for dealing with complex real-world problems where standard approaches, such as first principles modeling or explicit statistics modeling, are ineffective or unfeasible. CI typically works by displaying computational adaptivity and fault tolerance. Similarly, in numerical data, find IIoT node dependability and behavior patterns to identify similarities, grouping, correlations, and clustering.

Fig. 1 depicts the clustering process in the IoT network, at the starting of each round (which begins at a time), each sensor node chooses to be a CH with probability p so that the expected number of CH nodes for this round is p. Thus, if there are nodes in the network ensuring that all nodes are CHs, each node is required to be a CH once per round on average. This probability of becoming a CH is based on the assumption that all nodes start with an equal amount of energy and have data to send during each frame.

A. Network Model

IIoT can be modeled as an uni-directed graph $G = \{\nu, \varepsilon\}$, where ν represents the set of nodes that are interconnected by a set of edges ($\varepsilon \subseteq \nu \times \nu$) representing full-duplex wireless communication links. Here, ν and ε are changing over time due to the mobility of the nodes. Each node $n_i \subseteq \nu$, $l < i < N$ in the network is assigned a unique ID, where N is the number of network nodes. Two nodes n_i and n_j are said to be neighbors if they are within the transmission range of each other.

Algorithm 1 Allocating Cluster's in the IIoT Network

Input: E_{res}, $E_{\max}$, and C_{prob}.
Output: Probability of cluster (C_{prob}) and CH_{prob}
Procedure Clustering
begin:
Initialization
E_{res} = Residual energy in each node
$E_{\max}$ = Max energy in a node (with fully charged battery)
C_{prob} = initial percentage of CHs required in the network (5%)
Probability of becoming CH is assigned to all nodes using following equation
$CH_{prob} = C_{prob} \times \frac{E_{res}}{E_{\max}}$
CH selection
if node ($CH_{prob} < 1$) **then**
    node is tentative CH;
end if
if node ($CH_{prob} = 1$) **then**
    node is final CH;
end if
if node (is covered) **then**
    it joins cluster with CH to which it has low communication cost.
end if
if node (is not covered) **then**
    node elects itself as CH
end if
If a tentative CH node discovers a lower cost CH, it might become a normal node in a subsequent iteration.
If a node has a high E_{res} and a low communication cost, it might choose to become a CH at successive clustering intervals.
end Procedure

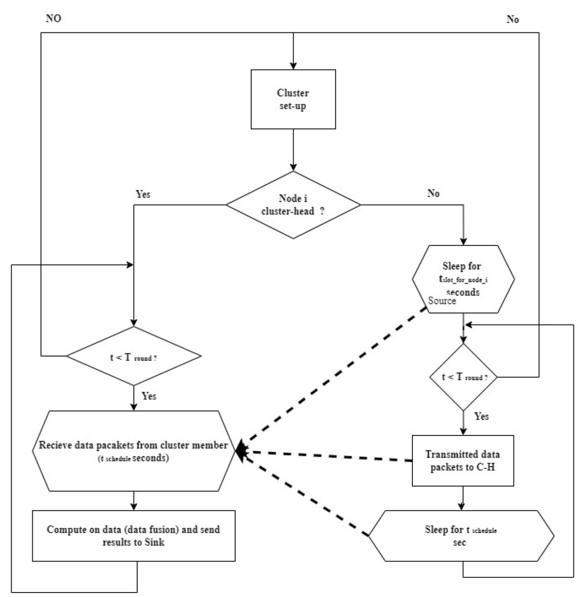


Fig. 1. Flowchart of the clustering process.

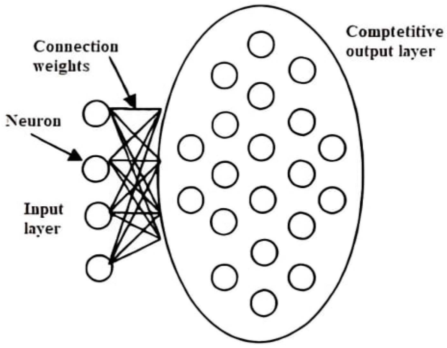


Fig. 2. Typical AOCA-NN model.

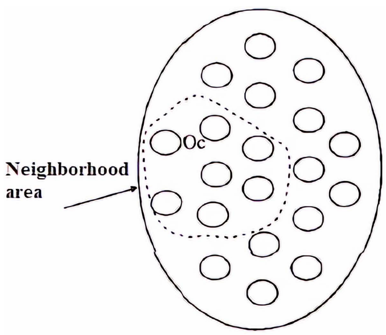


Fig. 3. Size of neighborhood winning output neuron ϱ_c.

An AOCA neural network (AOCA-NN) is made up of two layers: an input layer and a competitive layer at the output. Every neuron on the input layer is coupled to every neuron on the output layer, which is organized into 2-D grids. Fig. 2 depicts the number of input and output neurons in a typical AOCA-NN network.

The number of neurons in the AOCA network's input layer is proportional to the size of each input sequence. In this case, the input patterns are defined by the settings of different n_i's. For example, if the input parameter set size is n and each output parameter takes m binary bits, the total input neurons is obtained as per $| \nu | = mXn$. With an appropriate neighborhood value, the number of output neurons is set to be roughly double that of the input neurons. Each n_i neural network receives a collection of training input patterns. Assume the input pattern is $\eta = \{\eta_1, \eta_2, \ldots, \eta_{|\nu|}\}$.

Assume that there are $| \nu |$ neurons and τ neurons at the input/output layers. The output layer computes as per $\varrho = \{\varrho_1, \varrho_2, \ldots, \varrho_\tau\}$.

Each output neuron ϱ_j has $| \nu |$ receiving connections from the $| \nu |$ input neurons, where ω is the weight value assigned to each connection. As a result, the set of receiving connection weights for all the neuron ϱ_j at each output layer is $\omega = \{\omega_{j1}, \omega_{j2}, \ldots, \omega_{j|\nu|}\}$.

The input neuron pattern η is given at the input layer and neuron ϱ_j increases at the output layer based on the Canberra distance. The Canberra distance (D_j) is computed as follows, $D_j = \sqrt{\sum_{i=1}^{|\nu|} (\eta_i - \omega_{ji})^2}$.

At this point, the competitive output neuron with the shortest Canberra distance is closest to the current input pattern and is referred to as the winning neuron ϱ_c. Fig. 3 depicts an example of a neighborhood winning neuron ϱ_c. Equations (1) and (2) are used in the neighborhood size and weight update computation. The neighborhood λ size begins with a large enough size and lowers *wrt* learning iterations

$$\lambda_t = \lambda_o\left(1 - \frac{t}{T}\right) \tag{1}$$

where λ_t represents the actual neighborhood size, λ_o represents the beginning neighborhood size, t represents the current learning *epoch*, and T represents the total number of *epochs* to be completed. During the learning stage, the weights of the receiving connections of each winning neuron weights can be computed as per

$$\omega_{j_{\text{new}}} = \omega_{j_{\text{old}}} + \alpha\left(\eta - \omega_{j_{\text{old}}}\right) \tag{2}$$

where ω_j is the neighborhood of the winning output neuron ϱ_c and α is the learning rate parameter, and it has a typical value range of [0.2, ..., 0.5].

Competitive layers provide effective adaptive classifiers, but they have a few drawbacks. The first is that the choice of learning rate necessitates a tradeoff between learning speed and the stability of the final weight vectors. Slow learning occurs when the learning rate is close to zero. However, once a weight vector reaches the center of a cluster, it tends to stay there, and the learning rate reaches 1.0. However, once the weight vector reaches a cluster, it will continue to fluctuate between the different vectors in the cluster. The tradeoff between rapid learning and stability can be used for a high learning rate. The learning rate can be reduced to create stable prototype vectors.

B. CH Selection Using Fuzzy Logic

Fuzzy logic is used to choose CHs from among the n_is in a cluster. CH selection takes into account n_i characteristics like energy, storage, and speed of processing. Fuzzy rules are used as shown in Table I. Each sensor node is assigned a CH coefficient value based on the fuzzy rules used.

The generated membership function (μ) and their accompanying linguistic states represented in fuzzy logic are utilized to choose CH's from a group of n_i's. CH selection takes into account n_i factors, such as energy, storage, and speed of processing. The membership function for several n_i metrics from xbow mica2, such as energy, storage, and speed of processing as well as the CH co-efficient (Ch-coeff).

CH-coeff: Let "CH-coeff-threshold" be the CH coefficient threshold value for a n_i to become CH, where "CH-coeff" is computed by taking n_i properties like energy, storage, speed of processing, mobility. If the ch-coeff of node n_i > CH-coeff-threshold, then the CH-coeff value is high. If the n_i is considered as the CH; otherwise, n_i cannot be the CH. Continuous monitoring can be performed using n_i's with high energy, storage, and processing speed. Continuous monitoring can be performed using n_i's with the maximum energy, storage, and processing speed. Event monitoring can use n_i's with moderate energy, maximum storage, and high processing speed; critical monitoring can use n_i's with moderate energy and less energy and processing speed. n_i's with less energy, storage, or speed of processing is low cannot be used for monitoring.

TABLE I
FUZZY RULES FOR CH SELECTION

Rule number	CH-coeff	Energy in mw	Speed of processing in b/sec	Mobility factor (0 to 1)	Storage in kB
1	not used	less	Low	high	less
2	low	moderate	Medium	high	moderate
3	low	moderate	Medium	high	maximum
4	medium	moderate	High	medium	moderate
5	medium	moderate	High	medium	maximum
6	high	maximum	Medium	medium	maximum
7	high	maximum	High	low	moderate
8	high	maximum	High	low	maximum

These rules are used to calculate the CH-coeff for n_i values. In the first round of iteration, n_i with the highest CH coefficient CH-coeff value is chosen as CH, for the middle round of iteration, n_i with a medium CH-coeff value is chosen as CH, and at last in the final round of iteration, n_i with a low CH-coeff value is chosen as CH.

C. Clustering Using AOCA-NN

Clustering schemes divide n_is into tiny clusters. Each cluster has a CH, which collects data packets from its group of n_i's, aggregates the packets, and sends relevant data packets to the sink. If the node is within the sensing and transmitting range of other nearby nodes and the parameter of interest is well within the gradient value of deviation, that means sensed data is redundant. Clustering techniques are often used in IIoT, because of their efficient node coordination and leverage multihop routing between CHs to prevent long-distance transmissions.

The AOCA-NN is competitive, feedforward, unsupervised, and self-organizing capabilities. AOCA-NN can identify patterns in its input space during the learning process and automatically generates clusters to indicate different types of input classes; this computation is important in IIoT applications. The n_i parameters, such as storage, energy, the sensitivity of n_i's, and the network parameters, such as bandwidth, traffic, and interference are considered to cluster n_i's using AOCA-NN.

Number of Iteration: Let r is the number of iterations for which a n_i is considered as CH, PN is the number of n_i's called potential nodes that are eligible to become CHs in a cluster and X is the total number of transmission of data to the sink from a CH. The number of iteration r is computed as per

$$r = X \times \frac{i}{\mathrm{PN}}. \tag{3}$$

At the sink, using AOCA-NN and fuzzy logic clustering and Ch-coeff to each of the n_i is computed. Sink transmits this information as (cluster id, CH-coeff) to all n_i's. n_i with the highest coefficient value becomes the first CH. It informs this information to each n_i in its cluster. It receives data from n_i's, aggregates, and sends it to sink for r_1 number of iterations. The n_i's that have CH-coeff medium and high are considered as potential n_i's suitable as CH. After r_1 iterations, the second CH will take the role of the first CH, and the process repeats for r_2 number of iterations. This process repeats for r_3, r_4, and so on until all potential n_i's are used as CH's. Again, all the n_i's in the network send their network parameters to sink for new clustering and CH selection. The proposed AOCA-NN clustering routing protocol is described in Algorithm 2.

Algorithm 2 AOCA-NN Algorithm for Dynamic Clustering

```
Input: Let i = Number of iterations, I = Total number of iterations, r = Number of iterations CH undergoes; j = CH number; PN = Number of potential n_i's that can be used as CHs, y = random number that follows normal distribution.
Output: Forming of dynamic clusters and CH.
Procedure Dynamic clusters
begin:
In sink:
Receives: n_i's parameters and network parameters.
Clustering of n_i's are performed as per AOCA-NN.
'CH-coeff' is assigned for each n_i. using fuzzy logic with fuzzy based rule.
Transmit: (cluster id, CH-coeff) information to n_i's regarding n_i 'CH-coeff' and which cluster the n_i belongs to Within cluster:
for (j= 1 to n) do
    if i < j × r then
        for (k=l to r) do
            CH(j) receives data from n_is, aggregates and transmit to sink;
            i++;
        end for
    end if
end for
end Procedure
```

D. AOCA-HNN

Hop-field and Tank (1985) pioneered the use of HNN to tackle limited optimization problems. The objective function was represented as a quadratic energy function, and the related weights across neurons were calculated using gradient descent of the energy function. The energy function reduces monotonically as the number of repetitions increases, resulting in steady functioning. Hop-field has modeled the neural network as an analogue circuit that can solve multiobjective challenges. Every neuron is replicated electrically by a combination of nonlinear operational amplifiers, resistance, capacitance, and a current generator. The electrical circuit equivalents of a neural network and a neuron are depicted in Fig. 4. There are several feedback loops in the neural network, in that each neuron's output is sent back to the next neuron and does not give the self-feedback.

The ith neuron produces output voltage V_i, which is connected to input voltage U_i by the continuous learning function $g_i(u_i)$, as shown in Fig. 4. This learning function $g_i(u_i)$ is calculated using

$$V_i = g_i(u_i) = \frac{1}{1 + e^{-\theta_i u_i}} \tag{4}$$

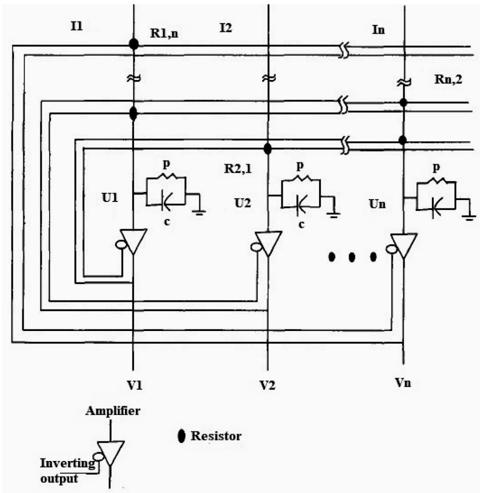


Fig. 4. AOCA-HNN.

where θ_i is the amplifier gain for the ith neuron. The output V_i is limited to the upper and lower saturation levels of 1 and 0, respectively. Kirchoff's current law is used to construct the circuit equation from the neuron model, as shown in

$$c_i \frac{dU_i}{dt} = \sum_{j=1}^{N} \frac{1}{r_{ij}} (V_j - U_i) - \frac{U_i}{p_i} + I_i. \tag{5}$$

The equation is rewritten as

$$c_i \frac{dU_i}{dt} = \sum_{j=1}^{N} T_{ij} V_j - \frac{U_i}{r_i} + I_i \tag{6}$$

where

$$\frac{1}{r_j} = \sum_{j} \frac{1}{r_{ij}} + \frac{1}{p_i} \tag{7}$$

and

$$T_{ij} = \frac{1}{r_{ij}}. \tag{8}$$

The T_{ij} in the preceding equation reflects the linking weights between the ith and jth neurons. The leakage resistance and capacitance of the amplifier are represented by r_i and c_i, respectively. In addition, each neuron gets an external bias input current I_i, which represents user-defined input to the neural network.

Equation (9) represents the Lyapunov energy function associated with the HNN

$$E = -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} T_{ij} V_{ij} V_j + \sum_{N}^{i=1} \frac{1}{r_i} \int_0^{V_i} g_i^{-1}(V)\ dV - \sum_{i=1}^{N} I_i V_i. \tag{9}$$

For symmetric network $T_{ij} = T_{ji}$, the time derivative of energy function E results as per

$$\frac{dE}{dt} = -\sum_{i=1}^{N} \frac{dV_i}{dt} \left(\sum_{j=1}^{N} T_{ij} V_j - \frac{U_i}{r_i} + I_i \right)$$

$$= -\sum_{i=1}^{N} \frac{dV_i}{dt} c_i \frac{dU_i}{dt}$$

$$= -\sum_{i=1}^{N} c_i \left(\frac{dV_i}{dt} \right)^2 \frac{dU_i}{V_i}. \tag{10}$$

TABLE II
SIMULATION PARAMETERS USED AOCA-NN

Simulation parameter	Values
Power considered in sensor node	CPU power, radio power, Total power
Sensor used	Temperature sensor
Threshold	I mV to 5 mV
Sampling rate	250ms to 1sec
Sensor node considered	Mica 2 mote
Number of nodes	100-1000
Operating system	TinyOS
Clustering	Location based
CH	Remaining power in the sensor node
Area of deployment	100 X 100 m

Because of the activation function increases monotonically, it is best to represent it using

$$\frac{dU_i}{dV_i} > 0. \tag{11}$$

With the substitution of (11) into (10), (10) results as

$$\frac{dE}{dt} \leq 0. \tag{12}$$

According to (12), energy E always declines (when $[dE/dt] \leq 0$) or remains constant ($[dE/dt] = 0$). The unchanging energy obtained when $(dV_i/dt) = 0$ indicated that a stable state had been established. As a result, the HNN always converges to a stable state. It is difficult to add the second term and compute according to

$$\sum_{N}^{i=1} \frac{1}{r_i} \int_0^{\infty} g_i^{-1}(V) dV. \tag{13}$$

The main difficulties with energy function may be eliminated. After ignoring the second component, the energy function (14) to be employed for optimization is as follows:

$$E = -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} T_{ij} V_{ij} V_j - \sum_{N}^{i=1} I_i V_i. \tag{14}$$

To simplify the dynamic model, all neurons are considered to have certain parameters r_i and c_i. After normalizing the bias current I_i and connectivity weights T_{ij} with capacitance c_i, (5) characterizing the dynamics of an HNN may be expressed as

$$\frac{dU_i}{dt} = \sum_{j=1}^{N} T_{ij} V_j - \frac{U_i}{r_i c_i} + I_i. \tag{15}$$

Combining equations (14) and (15), the time derivative (dU_i/dt) can be expressed as

$$\frac{dU_i}{dt} = -\frac{U_i}{r_i c_i} - \frac{\partial E}{\partial V_i}. \tag{16}$$

The minimum of energy function occurs at 2^N corners of N-dimensional hypercube defined by ($V \in 0, 1$).

TABLE III
COMPARISON TABLE OF AOCA WITH THE PEER ROUTING PROTOCOL

Cluster routing Protocols	Network	Cluster formation	CH selection	Algorithm complexity	CH role	Process dynamic	Location awareness
AOCA (Proposed)	Heterogeneous	Centralized	Node parameters and network dynamics	$o(n^2)$	Aggregating	Highly dynamic	Not required
ARFOR [23]	Homogeneous	Distributed	Random	$o(n^2)$	Relaying	Less	Required
MTCEE-LLN [24]	Homogeneous	Centralized	Random	$o(n^2)$	Relaying	Less	Required

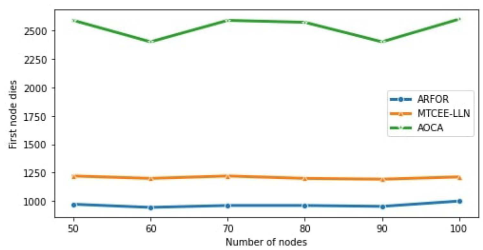


Fig. 5. FND in the network.

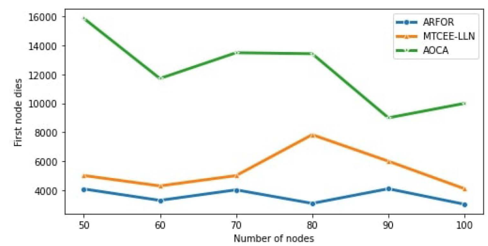


Fig. 6. LND in the network.

IV. RESULTS AND DISCUSSION

A. Simulation Setup

The TinyOS is used to establish an IIoT network with sensor nodes and random energy distribution among the nodes to create a realistic environment for IIoT using the TOSSIM. The simulation parameters are tabulated in Table II. Different parameters are configured for the proposed network using network components, interfaces, event handlers, and available sensors. In the proposed system, real-time event detection is implemented on Crossbow motes using a temperature sensor. The system is evaluated to assess its performance using characteristics, such as threshold values, cluster formation based on residual energy in nodes, event occurrence, and traffic flow in the system for synchronous and asynchronous monitoring.

B. Performance Evaluation Results

To evaluate the effectiveness of the proposed approach, simulation is performed based on IoT application characteristics, such as topology, reliability, movement direction, location, number of hops, energy levels, sensitivity, latency, and so on. For varying numbers of simulation parameters, the clustering of n_i using AOCA is estimated.

FND for n_i's in the network is as shown in Fig. 5. It can be seen that the first node death occurs for ARFOR and MTCEE-LLN in early compared to the proposed dynamic routing protocol. This result is useful in the analysis of the proposed routing protocol in comparison with other routing protocols, as FND occurs late and, hence, there is improvement in a network lifetime.

LND for n_i's in the network is as shown in Fig. 6. LND for the proposed routing protocol is occurring in a higher number of iterations. Hence, the network lifetime increases. Figs. 7 and 8 depict the number of packets sent by n_i to CH and CH to sink, respectively. It can be seen that for

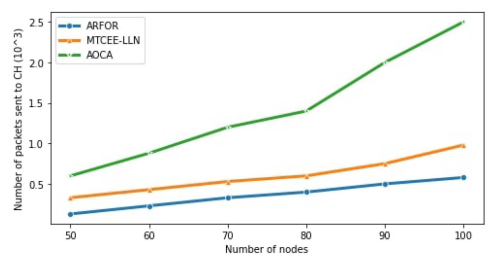


Fig. 7. Number of packets sent to CH.

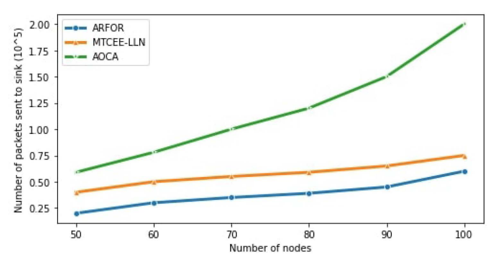


Fig. 8. Number of packets send to sink.

the proposed routing protocol AOCA, the number of packets sent is compared to ARFOR and MTCEE-LLN is shown in Table III. These results show that the throughput of the proposed AOCA protocol is more efficient compared with the peer routing protocol.

In the proposed routing protocol since clustering is made from n_i parameters, network dynamics, and applications requirement, clusters formed are highly dynamic and can be used for various applications requirement. Clusters and CH are selected based on resources available in the n_i, network parameters and application requirement is based on random selection. The rotation of CH, as well as the mobility of n_i's, are taken into account, energy consumption in n_i's is minimized, and the network is load balanced.

V. Conclusion and Future Scope

The proposed AOCA protocol employed neuro-fuzzy approaches for clustering and analysis to investigate cluster formation behavior based on the sensor node, network characteristics, and application requirements. This research work presented an improved Neuro-Fuzzy approach to improve self-managing intelligent systems that are self-organizing and self-sustaining, even while providing real-time computations and services in an ubiquitous networking environment. The proposed protocol used fuzzy sets for each cluster's dynamic clustering and resource management. The idea was to use CI to construct cluster routing and compute the optimal transmission path between source and destination. The suggested approach was combined with HNN. The future work of this article is that the dynamic node selection can be based on the Euclidean distances, such as the Manhattan distance, Chebyshev distance, etc. The higher level of dynamic clustering can be considered for sensor node behavior, such as the confidence level, security, trust, and reliability.

References

[1] J. Chase, "The evolution of the Internet of Things," Texas Instrum. Inc., Dallas, TX, USA, White Paper, 2013.

[2] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions," *Future Gener. Comput. Syst.*, vol. 29, no. 7, pp. 1645–1660, 2013.

[3] Q. Jing, A. V. Vasilakos, J. Wan, J. Lu, and D. Qiu, "Security of the Internet of Things: Perspectives and challenges," *Wireless Netw.*, vol. 20, no. 8, pp. 2481–2501, 2014.

[4] A. Ali, Y. Ming, T. Si, S. Iram, and S. Chakraborty, "Enhancement of RWSN lifetime via firework clustering algorithm validated by ANN," *Information*, vol. 9, no. 3, p. 60, 2018.

[5] J. A. Stankovic, "Research directions for the Internet of Things," *IEEE Internet Things J.*, vol. 1, no. 1, pp. 3–9, Feb. 2014.

[6] J. Duan, D. Gao, D. Yang, C. H. Foh, and H.-H. Chen, "An energy-aware trust derivation scheme with game theoretic approach in wireless sensor networks for IoT applications," *IEEE Internet Things J.*, vol. 1, no. 1, pp. 58–69, Feb. 2014.

[7] P. C. S. Rao, P. K. Jana, and H. Banka, "A particle swarm optimization based energy efficient cluster head selection algorithm for wireless sensor networks," *Wireless Netw.*, vol. 23, no. 7, pp. 2005–2020, 2017.

[8] A. Shankar, N. Jaisankar, M. S. Khan, R. Patan, and B. Balamurugan, "Hybrid model for security-aware cluster head selection in wireless sensor networks," *IET Wireless Sensor Syst.*, vol. 9, no. 2, pp. 68–76, 2019, doi: 10.1049/iet-wss.2018.5008.

[9] H. Farman et al., "Multi-criteria based zone head selection in Internet of Things based wireless sensor networks," *Future Gener. Comput. Syst.*, vol. 87, pp. 364–371, Oct. 2018.

[10] S. Kallam, R. B. Madda, C.-Y. Chen, R. Patan, and D. Cheelu, "Low energy aware communication process in IoT using the green computing approach," *IET Netw.*, vol. 7, no. 4, pp. 258–264, Jul. 2018.

[11] S. A. Latif et al., "Ai-empowered, blockchain and SDN integrated security architecture for IoT network of cyber physical systems," *Comput. Commun.*, vol. 181, pp. 274–283, Jan. 2022.

[12] C. Iwendi, S. U. Rehman, A. R. Javed, S. Khan, and G. Srivastava, "Sustainable security for the Internet of Things using artificial intelligence architectures," *ACM Trans. Internet Technol.*, vol. 21, no. 3, pp. 1–22, 2021.

[13] T. M. Behera, U. C. Samal, and S. K. Mohapatra, "Energy-efficient modified LEACH protocol for IoT application," *IET Wireless Sensor Syst.*, vol. 8, no. 5, pp. 223–228, Oct. 2018.

[14] D. Ngabo, D. Wang, C. Iwendi, J. H. Anajemba, L. A. Ajao, and C. Biamba, "Blockchain-based security mechanism for the medical data at Fog Computing Architecture of Internet of Things," *Electronics*, vol. 10, no. 17, p. 2110, 2021.

[15] Y.-K. Chiang, N.-C. Wang, and C.-H. Hsieh, "A cycle-based data aggregation scheme for grid-based wireless sensor networks," *Sensors*, vol. 14, no. 5, pp. 8447–8464, 2014.

[16] J. Xiao, Y. Tian, L. Xie, X. Jiang, and J. Huang, "A hybrid classification framework based on clustering," *IEEE Trans. Ind. Informat.*, vol. 16, no. 4, pp. 2177–2188, Apr. 2020.

[17] K. Zhan, C. Zhang, J. Guan, and J. Wang, "Graph learning for multiview clustering," *IEEE Trans. Cybern.*, vol. 48, no. 10, pp. 2887–2895, Oct. 2018.

[18] C. Lu, J. Feng, Z. Lin, T. Mei, and S. Yan, "Subspace clustering by block diagonal representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 2, pp. 487–501, Feb. 2019.

[19] S. Zhou et al., "Multiple Kernel clustering with neighbor-kernel subspace segmentation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 4, pp. 1351–1362, Apr. 2020.

[20] P. Chithaluru, F. Al-Turjman, T. Stephan, M. Kumar, and L. Mostarda, "Energy-efficient blockchain implementation for cognitive wireless communication networks (CWCNs)," *Energy Rep.*, vol. 7, pp. 8277–8286, Nov. 2021.

[21] J. H. Anajemba, C. Iwendi, M. Razzak, J. A. Ansere, and M. I. Okpalaoguchi, "A counter-eavesdropping technique for optimized privacy of wireless industrial IOT Communications.," *IEEE Trans. Ind. Informat.*, vol. 18, no. 9, pp. 6445–6454, Sep. 2022.

[22] S. A. Latif et al., "Ai-empowered, blockchain and SDN integrated security architecture for IoT network of cyber-physical systems," *Comput. Commun.*, vol. 181, pp. 274–283, Jan. 2022.

[23] P. Chithaluru, S. Kumar, A. Singh, A. Benslimane, and S. K. Jangir, "An energy-efficient routing scheduling based on fuzzy ranking scheme for Internet of Things," *IEEE Internet Things J.*, vol. 9, no. 10, pp. 7251–7260, May 2022.

[24] P. Chithaluru, F. Al-Turjman, M. Kumar, and T. Stephan, "Mtcee-LLN: Multilayer threshold cluster-based energy-efficient low-power and lossy networks for Industrial Internet of Things," *IEEE Internet Things J.*, vol. 9, no. 7, pp. 4940–4948, Apr. 2022.

Premkumar Chithaluru received the B.Tech. and M.Tech. degrees in computer science engineering from Jawaharlal Nehru Technological University Hyderabad, Hyderabad, India, in 2009 and 2011, respectively, and the Ph.D. degree in computer science engineering from the University of Petroleum and Energy Studies, Dehradun, India, in 2020.

He is currently working as an Associate Professor with the Department of Computer Science and Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad. His areas of research are WSN, VANET, Blockchain, IoT, IIoT, and AIoT.

Fadi Al-Turjman received the Ph.D. degree in computer science from Queen's University, Kingston, ON, Canada, in 2011.

He is a Full Professor and a Research Center Director of Near East University, Nicosia, Cyprus. He is also with the University of Kyrenia, Kyrenia, Turkey. His publication history spans over 350 SCI/E publications, in addition to numerous keynotes and plenary talks at flagship venues. He has authored and edited more than 40 books about cognition, security, and wireless sensor networks' deployments in smart IoT environments, which have been published by well-reputed publishers, such as Taylor and Francis, Elsevier, IET, and Springer. He is a leading authority in the areas of smart/intelligent IoT systems, wireless, and mobile networks' architectures, protocols, deployments, and performance evaluation in Artificial Intelligence of Things.

Prof. Al-Turjman has received several recognitions and best papers' awards at top international conferences. He also received the prestigious Best Research Paper Award from Elsevier *Computer Communications Journal* for the period 2015–2018, in addition to the Top Researcher Award for 2018 at Antalya Bilim University, Antalya, Turkey. He has led a number of international symposia and workshops in flagship communication society conferences. He currently serves as a Book Series Editor and the Lead Guest/Associate Editor for several top tier journals, including the IEEE Communications Surveys and Tutorials (IF 23.9) and the Elsevier *Sustainable Cities and Society* (IF 5.7), in addition to organizing international conferences and symposiums on the most up to date research topics in AI and IoT.

Manoj Kumar (Member, IEEE) received the M.Sc. degree in information security and digital forensics from Technological University Dublin (formerly, ITB Blanchardstown), Dublin, Ireland, in 2013, and the Ph.D. degree from The Northcap University, Gurugram, India, January 2019.

He has 12 years of research, teaching, and corporate experience. He is currently working on the post of an Associate Professor with the University of Wollongong, Dubai, UAE. He received fully funded scholarship for his M. Tech and M.Sc. programs from the Irish Government. He is mentoring six Ph.D. candidates and one M.Tech. by research candidate. He delivered two international research projects. He is currently involved in various research activities with high-impact researchers worldwide. He has published over 120 articles in international refereed journals and conferences. He published ten patents of which three are granted and two are in the examination stage. He published two textbooks and seven edited books. He edited eight journal special issues from the most famed publishers, such as Springer, Bentham Science, Taylor & Francis, and Elsevier. His specialization is cybersecurity, digital image processing using machine learning, IoT, digital forensics, networks, and information security.

Dr. Kumar has received the Best Researcher Award in 2020, an Outstanding Scientist Award in 2021, and a Young Researcher Award in 2021 from recognized bodies. He has delivered several guest lectures and seminars and chaired the session at various reputed international conferences. He is an associate editor, guest editor, and editorial board member of various journals of repute. He is serving as a Reviewer and an Editorial Board Member for many reputed journals, including the IEEE TRANSACTIONS (Q1 and Q2), and a technical program committee member for international conferences. He is also a member of numerous renowned bodies, including ACM, IAENG, ISTS, and UACEE.

Thompson Stephan (Member, IEEE) received the B.E. degree in computer science and engineering from Vel Tech Multi Tech Engineering College, Anna University, Chennai, India, in 2011, the M.E. degree in computer science and engineering from Easwari Engineering College, Anna University, Chennai, in 2013, and the Ph.D. degree in computer science and engineering from Pondicherry University, Puducherry, India, in 2018.

He is currently holding the position of an Assistant Professor with the Department of Computer Science and Engineering, Faculty of Engineering and Technology, M. S. Ramaiah University of Applied Sciences, Bengaluru, India. He has authored many technical research papers published in leading journals and conferences from IEEE, Elsevier, and Springer.

Dr. Stephan was awarded for his academic excellence in his master's degree with a university rank. He received the Best Researcher Award-2020, IEEE Bangalore Section, India. He is an active reviewer in many international reputed journals. He is currently the editor of two Springer books.

Sustainable Cities and Society 90 (2023) 104366

Contents lists available at ScienceDirect

Sustainable Cities and Society

journal homepage: www.elsevier.com/locate/scs

Energy-balanced neuro-fuzzy dynamic clustering scheme for green & sustainable IoT based smart cities

Premkumar Chithaluru [a], Fadi Al-Turjman [b,c], Manoj Kumar [d], Thompson Stephan [e,*]

[a] Department of Computer Science and Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad, India
[b] Research Center for AI and IoT, Faculty of Engineering, University of Kyrenia, Mersin 10, Turkey
[c] Artificial Intelligence Engineering Dept., AI and Robotics Institute, Near East University, Mersin 10, Turkey
[d] Faculty of Engineering and Information Sciences, University of Wollongong in Dubai, Dubai Knowledge Park, Dubai, United Arab Emirates
[e] Department of Computer Science and Engineering, Faculty of Engineering and Technology, M. S. Ramaiah University of Applied Sciences, Bangalore, Karnataka, India

ARTICLE INFO

Keywords:
IoT
QoS
Sustainable
Dynamic clustering
Neuro-fuzzy
Green communication protocols
Smart city

ABSTRACT

The Internet of Things (IoT) is a pervasive computing technology that provides solutions to critical sustainable smart city applications. Each sustainable application has its own set of requirements, including energy efficiency, Quality of Service (QoS), hardware, and software resources. Even though green IoT devices operate in a resource-constrained environment. Monitoring, recognizing, and responding to activities that entail continuous access to timely information in a partially or fully distributed ecosystem is a difficult task. To overcome the challenges of resource management in the IoT, we proposed an energy-efficient Dynamic Clustering Routing (DCR) protocol using a neuro-fuzzy technique for restricting the resources of IoT devices. The proposed protocol uses a dynamic self-organizing neural network to create dynamic clusters in a network. The test-bed analysis is for computing the real-time event detection and clustering sensor nodes using TinyOS. The simulation result shows that the proposed protocol achieved a significant gain over peer-competing well-known green communication routing protocols like Low-energy Adaptive Clustering Hierarchy (LEACH) and Low-energy Adaptive Clustering Hierarchy-Centralized (LEACH-C). The proposed model results show that using neuro-fuzzy logic is effective for sustainable IoT devices and green smart city applications in terms of resource management and dynamic clustering. The result analysis shows that the proposed protocol shows an average 35% significant gain on the First Node Dies (FND), Last Node Dies (LND), the number of packets sent to CH & BS, network convergence time, network overhead, and average packet delay to compare with the LEACH and LEACH-C.

1. Introduction

Recent advancements in wireless communication technology and advanced techniques, as well as the decrease in electrical and sensor equipment, have sparked great interest in green IoT research (Abbasi & Shahid Khan, 2018; Javed et al., 2022). A Sustainable IoT network is made up of several sensor nodes that may communicate wirelessly and is widely placed inside a designated region of green smart cities (Chithaluru, Tiwari, & Kumar, 2021f; Mukherjee, Goswami, Yang, Yan, & Daneshmand, 2020). They use wireless channels to send and receive data from other nodes. The sensor nodes must be equipped with the ability to recognize nearby neighbors and contribute to the development of wireless networks using green communication protocols (Chithaluru, Khan, Kumar, & Stephan, 2021c; Saleem, Afzal, Ateeq, Kim, & Zikria, 2020).

Sustainable IoTs are application-specific, varying as per smart city applications. Small node size, low node cost, power efficiency, durability, extensible, self-configure, a place that enables, high availability, flexibility, safety, manufacturing costs, an environment in which it operates, sensor topology, hardware restrictions, and communication channels of all IoT networks (Jain et al., 2022; Zikria, Kim, Hahm, Afzal, & Aalsalem, 2019). These elements are significant because they serve as a framework for developing a protocol or algorithm for sensor networks. IoT efficiency may be increased by the unified network (Chithaluru, Al-Turjman, Kumar, & Stephan, 2020a), targeted, and resource strategies (Zikria, Yu, Afzal, Rehmani, & Hahm, 2018). In this case, the sensor nodes have natural capabilities for detecting nearby neighbors and assisting in the development of a wireless

* Corresponding author.
E-mail addresses: bharathkumar30@gmail.com (P. Chithaluru), fadi.alturjman@neu.edu.tr (F. Al-Turjman), wss.manojkumar@gmail.com (M. Kumar), thompsoncse@gmail.com (T. Stephan).

https://doi.org/10.1016/j.scs.2022.104366
Received 17 April 2022; Received in revised form 14 November 2022; Accepted 17 December 2022
Available online 16 January 2023

network using a existing routing protocols (Chithaluru, Tiwari, & Kumar, 2019), which can improve network activities such as network administration, sensor data handling, and aggregating. As a result, there are several benefits to clustering and their evaluation depending on system factors and technical specifications (Chithaluru, Tiwari, & Kumar, 2021e; Manman, Xin, Goswami, Mukherjee, & Yang, 2020). The complex challenge will be systematic clustering by selecting important parameters and responding to their dynamic characteristics in bandwidth-restricted wireless networks (Chithaluru & Prakash, 2020).

IoTs are dynamic, with greater uncertainty associated with the input traffic and other environmental parameters. They are subjected to unexpected overloads and failures. Neuro-fuzzy logic appears to be a promising approach to address some of the important decision-making aspects of the IoT. Since fuzzy sets provide a robust mathematical framework for dealing with real-world imprecision and non-statistical uncertainty, the application of neuro-fuzzy logic in communication networks is recent and relatively less extensive than in automatic control. The classical set theory allows elements to be either included in a set or excluded from it. This is in contrast with human reasoning, which includes a measure of imprecision or uncertainty, marked by the use of linguistic variables such as self-sustaining, self-organizing, etc. This approximate reasoning is modeled by fuzzy logic, which is a multi-valued logic that allows intermediate values to be defined between conventional threshold values. Fuzzy systems allow the use of fuzzy sets to draw conclusions and make decisions. Fuzzy sets differ from classical sets in that they allow an object to be a partial member of a set. Self-sustaining and self-organizing are the techniques used in the IoT to prolong the network's lifetime. Thus, the neuro-fuzzy approach plays an important role in increasing network lifetime, data delivery, and data aggregation.

In this paper, the authors developed a dynamic clustering approach and its analysis to investigate the dynamic characteristics of the clustering process depending on system characteristics and technical specifications. The scheme entails using cognitive computing to construct dynamic clustering based on sustainable IoT characteristics and green smart city applications.

1.1. Dynamic clustering in green smart cities

Sustainable IoT is used to sense the required parameter of interest in a variety of green applications (Al-Mutiri, Al-Rodhaan, & Tian, 2018). These smart city applications have their requirements based on sensor parameters and simulation parameters used in the network as shown in Fig. 1. For example, real-time applications like military, disaster management, and telemedicine require sensor nodes with high resources such as a battery, storage, efficient processing, bandwidth, etc., and for contiguous monitoring like environmental monitoring, habitat monitoring, precision agriculture applications, sensor nodes having fewer resources like less power, low processing speed, are enough for the normal operation (Chithaluru, Singh, & Sharma, 2020b). These examples show that sustainable IoT is used for different green applications, and its efficient routing demands in terms of network lifetime, data packet delivery, and delay. Clustering is one of the techniques used to increase network lifetime in sustainable IoT. Also, IoT is used for green applications in terms of clustering schemes to show the significant gain on each QoS. The illustration shows that sustainable IoT requires dynamic clustering that meets the requirement of various green applications (Ramakuri, Chithaluru, & Kumar, 2019).

1.2. Computational intelligence

Computational Intelligence (CI) is a novel and potential approach to solving problems of real-time, optimal, and adaptive solutions (Jena, Ammoun, & Chithaluru, 2022). CI is considered a research initiative for neural network models, fuzzy rules, and evolutionary computation. Artificial neural networks (ANNs) have emerged as powerful and

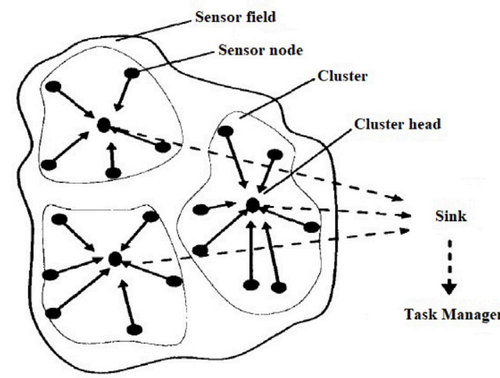


Fig. 1. Sustainable IoT network.

widespread data structures capable of significant learning. Some of the computations that are performed by neural networks are classification, regression, and constrained function optimization. The advancements in ANN technology may bring features ideally suited to solving some difficult challenges in sustainable IoT. IoTs are complex and dynamic, with a high degree of uncertainty related to input throughput and other environmental characteristics. They are prone to unanticipated load variations and outages. Fuzzy looks to be a potential solution for addressing such critical elements of sustainable IoTs (Chithaluru, Stephan, Kumar, & Nayyar, 2022). It provides a strong mathematical foundation for dealing with a real-world lack of precision and non-statistical ambiguity. According to a comprehensive investigation and review of the literature, extant research on fuzzy inference systems in network technologies extends from scheduling, buffer planning, and load control to forwarding, available bandwidth, infrastructure management, and quantified network performance evaluation.

1.3. Research challenges in sustainable IoT

IoTs are becoming more pervasive in the existing system and emerging applications are mentioned above. Hence there is a need to study and understand the system behavior in a different environment based on the potential application requirements and issues, including resource management, self-organization, routing, mobility, scalability, security, and data aggregation. Techniques should provide the solution for most of the above issues. Along with the issues and challenges in IoTs, the techniques should address application requirements such as real-time event detection, location-based monitoring, operating environment decisions, and so on.

1.4. Limitations of neuro-fuzzy approach

The neuro-fuzzy approach has significant limitations, such as dimensionality problems and learning rate variability, that further restrict implementations to large dataset challenges. Because of its complex structure and gradient learning, neuro-fuzzy has a high computational cost. This is a substantial constraint for large-input applications. The disadvantages are defined broadly as,

- Difficulty in interpreting functionality.
- Complexity in determining the number of layers and neurons.
- Not reliable in the network topology changes, which would require improvements to the fuzzy rules.
- Nature and number of decision variables.
- Scope of a membership function.
- Curse of dimensional space. Furthermore, the trade-off between interpretability and accuracy is considered a critical issue.

1.5. Resource management in sustainable IoT

The applications of IoTs are diverse, and they are deployed in completely different environments, each with a unique set of requirements. For example, some applications would need real-time data delivery, while others may need secure and reliable data delivery (Chien, 2007). This has led to several IoT protocols, proposed over the years, designed to address a specific set of application requirements. For example, IoT for disaster prevention, such as the detection of forest fires, tsunamis, volcanoes, etc., requires real-time data delivery and a delay guarantee as to the most critical requirements. On the other hand, if we are designing IoTs for military or security applications, sender authentication, and secure data transmission are more important. Since IoT is used for a wide variety of applications, research on IoT is necessary so that IoT technology can be used ubiquitously anywhere, anytime.

The research paper deals with the resource management issue of sustainable IoTs and it involves developing a model for tackling resource management challenges using dynamic clustering, developing an analytical model, designing and developing algorithms in IoT, and then constructing a simulation environment to evaluate the performance of the proposed technique and the model. Here, dynamic clustering is considered using neuro-fuzzy concepts and theories to manage resource management in sustainable IoT.

1.6. Problem statement

During the network planning and dimensioning phase, determining the optimal deployment of sensors to minimize energy consumption and prolong the network lifetime becomes an important problem. As a result, most existing peer protocols, as well as future protocols, must design and develop a methodology to analyze energy consumption, data aggregation, degree of confidence, network lifetime, and sensor networks in a given phenomenon of interest. IoT is used for various network-related applications, including environmental monitoring, telemedicine, and military field surveillance. Due to such a wide variety of applications, the sensor network needs a different and sometimes unique set of requirements. Developing a routing protocol for a network with resource constraints for remote application requirements is a challenging issue. Clustering is one of the techniques used for energy management in the IoT. Thus, we proposed a dynamic cluster routing protocol for IoT. The technique involves the sensor nodes in clustering depending on QoS parameters, network dynamics, and application requirements.

1.7. Research objectives

IoTs have proven to be extremely effective in a wide range of green smart city applications, resolving critical and often life-threatening issues. In this research work, neuro-fuzzy logic is used for dynamic clustering in the IoT network.

- The application of the neuro-fuzzy approach to the IoT results in self-sustaining, self-organizing, and conscious cognitive technologies that provide authentic computing and services in a ubiquitous communication network.
- Even most IoT problems are real-time in nature and necessitate fast computation, real-time optimal solutions, and being adaptive to the situation of events and data traffic in order to achieve desired results. Hence neural networks and fuzzy sets would be appropriate candidates for implementing most of the computations involved in the issues of resource management in sensor networks.
- A model is proposed for sustainable IoT to detect real-time events and nodes in a cluster to prolong the lifetime of the network.
- The proposed work aims to provide a dynamic clustering technique for resource management using the Neuro-Fuzzy technique. The approach entails using CI to create clustered routing mechanisms. The clustering of nodes is dependent on application characteristics, network functions, and technical specifications. The approach entails evaluating clustering and behavior for available resources and client-required resources.

The major contributions of the paper are as follows,

- Proposed an energy-efficient Dynamic Clustering Routing (DCR) protocol using a neuro-fuzzy technique for restricting the resources of IoT devices.
- Proposed protocol used the dynamic self-organizing neural network for creating the dynamic clusters in a network.
- The simulation result shows that the proposed protocol achieved a significant gain over peer-competing well-known green communication routing protocols like Low-energy Adaptive Clustering Hierarchy (LEACH) and Low-energy Adaptive Clustering Hierarchy - Centralized (LEACH-C). The proposed model results show that using neuro-fuzzy logic is effective for sustainable IoT devices and smart city applications in terms of resource management and dynamic clustering.

The remainder of the paper is organized as follows, various peer-competing routing protocols used in sustainable IoT are discussed in Section 2. The proposed dynamic clustering model is discussed in Section 3. The simulation and results are discussed in Section 4. Finally, Section 5 provides a conclusion and directions for future work.

2. Related work

In this section, various peer-competing approaches that support for clustering scheme are discussed.

Joshi et al. (2022), to arrange the sensors in an IoT into clusters, the authors suggested a decentralized, randomized clustering technique. The technique constructs a network of CHs, and it was discovered that the number of layers in the network increases energy efficiency. Stochastic geometry is employed to develop solutions for the values of the algorithm's parameters that reduce the total power consumption in the network when all sensors transmit information to the sink via the CHs (Tanwar, Balamurugan, Saini, Bharti, & Chithaluru, 2022).

Joshi et al. (2022a), developed a physical, Medium Access Control (MAC) network cross-layer analytical technique for calculating the ideal number of clusters in a high-density sensor network to reduce energy usage. Many effects may be incorporated into the cross-layer design, such as log-normal shadowing and a two-slope route loss model in the physical layer, as well as different MAC scheduling and multi-hop routing algorithms.

Chauhan, Chandra, and Maheshkar (2016), proposed a decentralized approach for clustering an ad-hoc sensor network. Each sensor decides whether to build a new cluster or use a random waiting timer and local node parameters. The technique works without a proactive operation and does not require the location of the nodes in advance.

Wang et al. (2017), presented a sensor node inside a cluster that assesses its relative energy usage in comparison to other nodes in the same cluster. Sensor nodes autonomously choose a time frame in which they will operate as a CH in the next round based on the proportionate quantity of energy consumed in the current round.

Chithaluru, Al-Turjman, Kumar, and Stephan (2021a), suggested a top-down cluster and cluster-tree construction process that enables successful message delivery to a sink as well as within the network nodes (Musaddiq et al., 2018).

Chithaluru, Kumar, Singh, Benslimane, and Jangir (2021d), developed a combined clustering and prediction approach. The paper discusses an energy-efficient clustering and integrating dynamic enabling/disabling prediction technique that employs a sleep/awake schedule for query processing.

Table 1
Performance comparison of proposed Vs. peer existing protocols.

Protocol	Self-organization	Self-sustaining	CH election	Mobility	Communication Inter cluster	Communication BS and CH
LEACH	Yes	Limited	Threshold	No	Single-hop	Single-hop
LEACH-C	Yes	Good	Residual energy	No	Single-hop	Single-hop
E-LEACH	Yes	Limited	Residual energy	No	Single-hop	Single-hop
M-LEACH	Yes	Good	Threshold	No	Single-hop	Multi-hop
V-LEACH	Yes	Limited	Residual energyand distance	No	Single-hop	Single-hop
Proposed	Yes	Very Good	Fuzzy inference based	Yes	Single-hop	Multi-hop

Chithaluru, Prakash, and Srivastava (2018), studied the event-driven clustered IoT, which is a probabilistic technique for analyzing network lifespan when events occur randomly over the network field. The sensor's packet transmission rate is modeled using the principles of coverage processes and Voronoi complex shapes. The likelihood of individual sensors attaining a specific lifespan is then calculated. This likelihood is then used in the investigation of cluster longevity.

Chithaluru and Prakash (2018), suggested a distributed energy-efficient clustering technique for heterogeneous IoTs that is randomized and controlled. The protocol divides the network into dynamic clusters. It proposes a balanced and dynamic technique for calculating the likelihood of a CH election (Hasrouny, Samhat, Bassil, & Laouiti, 2017).

He and Zhu (2012), proposed a low streaming data delivery methodology for cluster-based IoTs with mobile sinks. It ensures the end-to-end connection between the source and the mobile sinks by implementing a cross-cluster handover mechanism and a path redirection system while eliminating the continual transmission of the mobile sink position as it moves across different clusters (Chithaluru, Al-Turjman, Stephan, Kumar, & Mostarda, 2021b). Table 1 shows that performance comparison proposed Vs. peer existing protocols.

2.1. Research gaps identified in literature

Various limitations identified in the literature as follows,

- Several cluster-based routing algorithms are only effective in small areas or with a limited node density.
- Few Cluster-based routing protocols are appropriate for fixed cluster formation.
- The distribution of Cluster Heads (CHs) is focused on one region only in specific cluster-based routing protocols (LEACH and LEACH-C).
- For time-critical applications, cluster-based routing methodologies are inefficient (LEACH).
- A few cluster-based routing protocols allow all CHs to transfer collected data to a sink, causing more energy loss in the network.
- Several cluster-based routing schemes communicate information across the network using prediction-based methods that ignore remaining energy in the network nodes, which causes the early death of CHs.

3. Proposed method

Clustering is the preferred design for traffic-free nodes that would communicate data packets to near CHs within a predetermined radius of the network. As a result, the maximum number of packets is communicated to the sink and node alive in the network for a longer time. A CH will be present in each cluster and CH collect information from their group nodes and then transfer it to a distant sink, involving high energy transmission, data gathering, and relaying pertinent data packets to the sink. It reduces energy usage by not allowing all nodes to process, extending the network's lifespan. IoT necessitates dynamic clusters such as a group of nodes with specific energy, a group of nodes with high throughput, a group of nodes with wide bandwidth, or a combination of node properties. Dynamic clusters and their applications are shown in Fig. 2.

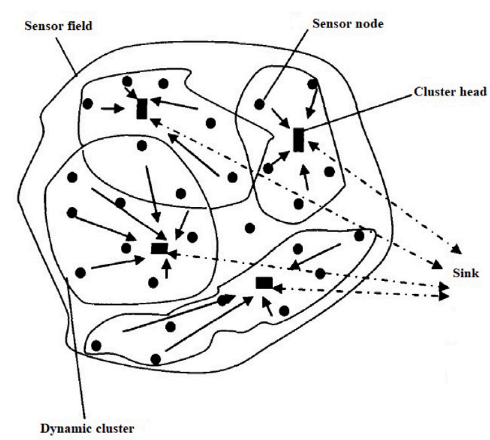


Fig. 2. Dynamic clusters in IoT.

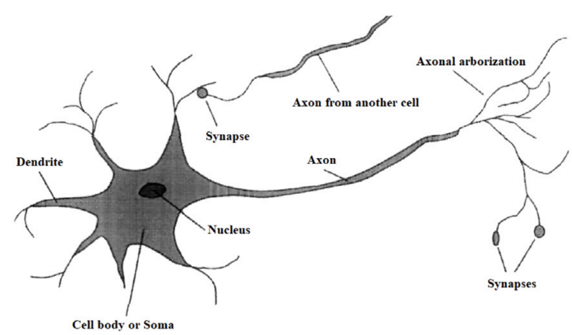


Fig. 3. Biological neuron.

3.1. Neural-fuzzy logic

The human brain is made up of a vast number of neuron cells that process information, more than a billion in total. Each cell functions similarly to a basic processor. The vast interaction of all cells and their concurrent processing only allows the brain's powers to be realized. Fig. 3 depicts a real neuron. Neurons are branching fibers that emerge from the neuron, also known as the soma. A neuron's soma, or neuron, includes the nuclei and other components that facilitate chemical analysis and neurotransmitter synthesis. An axon is a single fiber that transport signals from the soma to the connecting locations of other neurons, muscle, or organs. Axon Hillock is the summation point for collecting reports. The combined impact of all cells that respond to incidents to a specific cell at any particular time determines whether or not an electrical impulse is triggered at the axon terminal and transmitted up the axon. A synapse is a place at which neurons or a cell and muscles or organs communicate. At some of these connections, electromechanical communication between neurons occurs. Fig. 4 depicts a basic NN model. A neuron is made up of three essential features weights, a cutoff, and a training algorithm.

Weight factors W: The values $W_1, W_2, \dots, W_n$ are the weights with weight vector as, $W^T = [W_1, W_2, \dots, W_n]^T$ of input vector $X =$

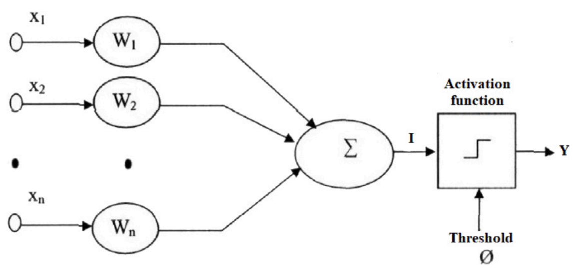


Fig. 4. Neuron model used in proposed.

$[x_1, x_2, \dots, x_n]$. Each input is combined with the corresponding weight of the neuron network $X.W^T$ to determine the total weight of the training dataset $X = [x_1, x_2, \dots, x_n]$. The $+ve$ weight escapes the output node, whereas the $-ve$ weight blocks it. We get $I = X.W^T$ and further its computed as per Eq. (1).

$$I = [x_1, x_2, \dots, x_n].[w_1, w_2, \dots, w_n] = (x_1 w_1 + x_2 w_2 + \dots + x_n w_n) = XW^T \quad (1)$$

3.2. Architecture of proposed neuro-fuzzy network

NNs are balanced ordered networks with nodes representing artificial neurons and directed weights indicating relationships between output units and input. There are two types of neural networks based on their connection pattern (architecture).

- Feed-forward (FF) NNs
- Feedback NNs

FF-NNs do not have loops, but response NNs do because of feedback links. These neurons of the most common type of FF-NN, known as multi-layer activation functions, are organized into layers with continuous interconnections. Different connection levels lead to different connections. FF-NN, in general, is stable, providing only one set of target values instead of a series of results from a single input. To the extent that their response to an input is autonomous of the previous state of the network, FF-NN networks have no storage.

CI has been used in the suggested clustering approach. CI is a collection of environment computing techniques and approaches for addressing complicated actual problems where traditional practices, such as first principles simulation or explicit statistics modeling, are inefficient or impossible. Typically, CI works with statistical information and does not employ information in the AI paradigm, exhibiting computing flexibility with high availability. In a similar way, the numerical data to identify IoT's node reliability and behavior patterns to identify similarities, grouping, correlations, and clustering.

3.3. Network model

IoT can be modeled as a uni-directed graph $G = (V, E)$, where V represents the set of nodes that are interconnected by a set of edges $E \subseteq (V \times V)$ representing full-duplex wireless communication links. Here V and E are changing over time due to the mobility of the nodes. Each node $n_i \subseteq V, l \leq i \leq N$ in the network is assigned a unique ID, where N is the number of nodes in the network. Two nodes n_i and n_j are said to be neighbors if they are within the transmission range of each other.

The suggested scheme collects characteristic features or parameters of each node, such as storage, the energy available in the nodes, hops from source to CH, and node awareness, for strategic planning at the sink, where classification evaluation is managed to carry out using the Gradient-based Feed-Forward Neural Network (GFF-NN) automated system to scatter the nodes based on the measurements of significance.

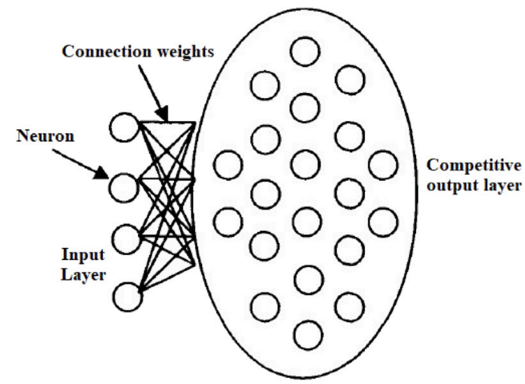


Fig. 5. Structure of GFF-NN.

GFF-NN is made of two parts: a sensing unit and an output competing layer. Each input layer communicates with every neuron on the output units, which are organized in two-dimensional blocks. Fig. 5 depicts a structure of GFF-NN with input and output layers.

The input data of a GFF-NN contains the same number of layers as the dimension of each input sequence. The input patterns are defined by the configuration of multiple nodes. If the feature set size is n and each variable needs m binary data (Binary NN), the number of information layers is $|\ V\ | = mXn$. The layers on the output units arrange the connections among the input vectors identified by the competing layers and may construct a topological map from an initial random position, with the final map showing the fundamental connections among the sequences defined. The number of neurons in each layer (competitive layered neurons) is governed by the number of distinct configurations (categories) in the training sets. With an appropriate neighborhood value, the number of available neurons is set to almost double the size of the layers of neurons. The network is fed a sequence of learning input sequences matching each node. The input sequence will be $I = (I_1, I_2, \dots, I_{|V|})$.

When the input layer is comprised of artificial neurons with $|\ V\ |$ values. Suppose the output (competing) layer has Q layers, with O_j indicating the jth output layer. As a consequence, the output unit will be $O = (O_1, O_2, \dots, O_Q)$.

Each output neuron O_j has V connection requests from the $|\ V\ |$ input neurons. The weight value applied to each connection is denoted by W. So, for each output Oj on the output units, the set of inbound network parameters is $W_j = (W_{j1}, W_{j2}, \dots, W_{j|V|})$.

When an input sequence I is provided in the neural network, the Euclid distance measure D_j of an unit O_j in the output nodes is $D_j = \sqrt{\sum_{i=1}^{|V|}(I_i - W_{ji})}$.

At this point, the competing output unit with the shortest distance function is nearest to the present input sequence, and the same is referred to as the winning neuron O_c. Fig. 6 depicts an example of a neighborhood centered on a winning neuron O_c. Eq. (2) computes the neighborhood size and weight update. The size of the neighborhood h begins large enough and shrinks with training repetitions, i.e.,

$$h_t = h_0(1 - \frac{t}{T}) \quad (2)$$

Where, h_t represents the actual neighborhood size, h_o represents the beginning neighborhood size, t represents the current learning $epoch$, and T represents the total number of $epochs$ to be completed. In this case, a $epoch$ is when the channel has once passed through all of the nodes' training sets. During the stage of learning, the weights of each competing neuron's incoming connections in the winner neuron's neighborhood, as well as the winning neuron, are adjusted. The total weights are updated according to Eq. (3),

$$W_j new = W_j old + \alpha(I - W_j old) \quad (3)$$

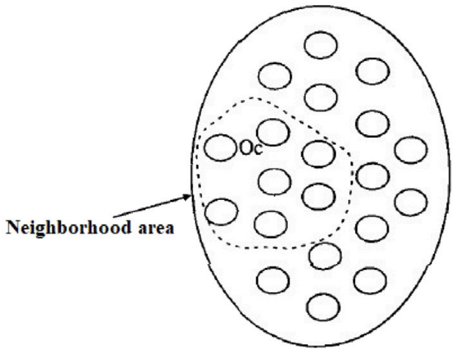


Fig. 6. O_c - the size of the neighborhood impact around a neuron.

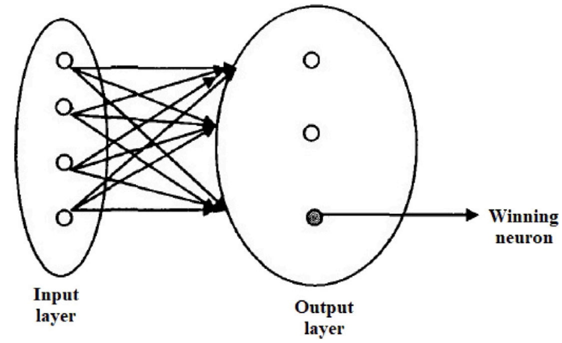


Fig. 7. Structure of the GFF-NN.

Table 2
Acronyms.

Symbols	Meaning
W	Weights
X	Training data set
G	Uni-directed graph
E	Edges in the graph
V	Vertices in the graph
I	Input sequences in the network
O	Output unit sequence in the network
D	Distance between the nodes
h	Actual neighborhood size
t	Current learning *epoch*
T	Total number of *epoch*
O_j	Close to the winning node O_c
α	Proportional gain

Table 3
Sensor node parameter range of Mica2 mote.

Parameters	Value
Battery level	0 to 3 mWatts
Storage	0 to 128 KB
Transmission Speed	0 to 7.3 MHz
Hops	1 to 3

O_j is close to the winning node O_c. Where α is the proportional gain variable, and its normal recommended value is [0.2, …., 0.5].

3.4. Example of dynamic clustering

In this section, the clustering of nodes using GFF-NN is illustrated. Sensor node parameters of Mica2 motes from CrossBow such as a battery, storage, mobility, processing limits, hops from the sink, bandwidth, etc., can be considered for clustering of sensor nodes in IoT. Table 3, Table 4, and Fig. 7 show the sensor node parameter range of Mica2 mote, the sensor node parameter, and the structure of the neural network considered for clustering respectively. If a sensor node has sufficient resources then use the vector coefficient as 1. otherwise as 0. A Kohonen self-organizing pattern is being used to aggregate four nodes with dimensions (1, 1, 1, 1); (0, 0, 0, 1); (1, 1, 0, 0); (1, 1, 0, 0). (0, 0, 1, 1). As shown in Fig. 7, the highest number of nodes that may be constructed is m=3. Assume the training set (periodic diminishing) is $\alpha(0) = 0.3$ and $\alpha(t+1) = 0.2\alpha(t)$. With just three sets available and only one cluster's parameters adjusted at every stage (i.e., N_c=0), determine the set of weights using only one *epoch* of learning. Table 2 referring for acronyms and variables.

Step 1: The initial weight matrix is:

$$\begin{bmatrix} 0.2 & 0.4 & 0.1 \\ 0.3 & 0.2 & 0.2 \\ 0.5 & 0.2 & 0.5 \\ 0.1 & 0.1 & 0.1 \end{bmatrix}$$

The radius at initial: N_c=0; First learning rate: $\alpha(0) = 0.3$.
Step 2: Start the training
Step 2 (1): Steps 3–5 apply to the third input vector x_1=(1,1,1,0).
Step 3:

$$I(1) = (1-0.2)^2 + (1-0.3)^2 + (1-0.5)^2 + (0-0.1)^2 = 1.39$$
$$I(2) = (1-0.4)^2 + (1-0.2)^2 + (1-0.3)^2 + (0-0.1)^2 = 1.5$$
$$I(3) = (1-0.1)^2 + (1-0.2)^2 + (1-0.5)^2 + (0-0.1)^2 = 1.71$$

Step 4: The training data is the one that is nearest to output node 1. As a result, node 1 is the winner. Node 1's frequency should be changed.
Step 5: The winning unit's weights have been updated:

$$W_1 new = W_1 old + \alpha(x - W_1 old)$$

Step 2 (2): Steps 3–5 apply to the third input vector x_2=(0,0,0,1).

$$\begin{bmatrix} 0.44 & 0.4 & 0.1 \\ 0.51 & 0.2 & 0.2 \\ 0.65 & 0.3 & 0.5 \\ 0.37 & 0.1 & 0.1 \end{bmatrix}$$

Step 3:

$$I(1) = (0-0.44)^2 + (0-0.51)^2 + (0-0.65)^2 + (1-0.37)^2 = 1.2731$$
$$I(2) = (0-0.4)^2 + (0-0.2)^2 + (0-0.3)^2 + (1-0.1)^2 = 1.1$$
$$I(3) = (0-0.1)^2 + (0-0.2)^2 + (0-0.5)^2 + (1-0.1)^2 = 1.11$$

Step 4: The training data is the one that is nearest to output node 2. As a result, node 2 is the winner. Node 2's frequency should be changed.
Step 5: The winning unit's weights have been updated:

$$W_2 new = W_2 old + \alpha(x - W_2 old)$$

Step 2 (3): Steps 3–5 apply to the third input vector x_3=(1,1,0,1).

$$\begin{bmatrix} 0.44 & 0.28 & 0.1 \\ 0.51 & 0.14 & 0.2 \\ 0.65 & 0.21 & 0.5 \\ 0.37 & 0.37 & 0.1 \end{bmatrix}$$

Step 3:

$$I(1) = (1-0.44)^2 + (1-0.51)^2 + (0-0.65)^2 + (1-0.37)^2 = 1.1131$$
$$I(2) = (1-0.28)^2 + (1-0.14)^2 + (0-0.21)^2 + (1-0.37)^2 = 1.439$$
$$I(3) = (1-0.1)^2 + (1-0.2)^2 + (0-0.5)^2 + (1-0.1)^2 = 1.71$$

Step 4: The training data is the one that is nearest to output node 1. As a result, node 1 is the winner. Node 1's frequency should be changed.
Step 5: The winning unit's weights have been updated:

$$W_1 new = W_1 old + \alpha(x - W_1 old)$$

$$\begin{bmatrix} 0.608 & 0.28 & 0.1 \\ 0.657 & 0.14 & 0.2 \\ 0.455 & 0.21 & 0.5 \\ 0.259 & 0.37 & 0.1 \end{bmatrix}$$

Table 4
Sensor node parameters considered in the example.

Node ID	Parameters considered				
	Battery (mW)	Storage (KB)	Bandwidth (Hz)	Transmission Speed	Hops from sink
1	1	1	1	0	1
2	1	0	1	1	0
3	1	0	0	1	1
4	1	1	1	1	1

Table 5
Simulation parameters used GFF-NN.

Simulation parameter	Values
Power considered in sensor node	CPU power, radio power, Total power
Sensor used	Temperature sensor
Threshold	I mV to 5 mV
Sampling rate	250ms to 1sec
Sensor node considered	Mica 2 mote
Number of nodes	100-1000
Opertaing system	TinyOS
Clustering	Location based
CH	Remaining power in the sensor node
Area of deployment	100 X 100 m

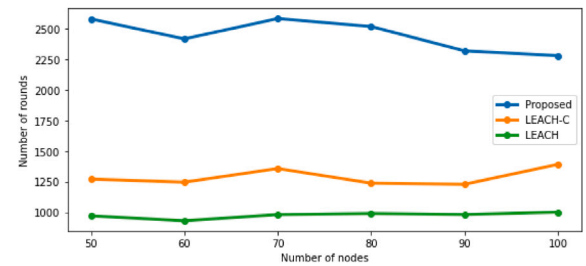


Fig. 8. FND in the network.

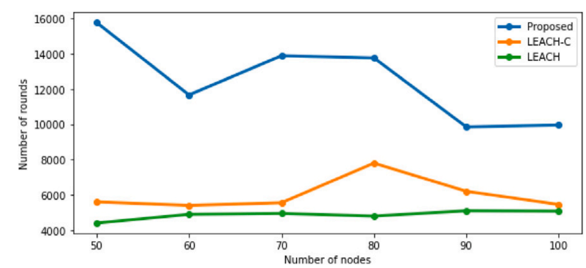


Fig. 9. LND in the network.

Step 2(4): Steps 3–5 apply to the fourth input vector x_4=(0,0,1,1).
Step 3:

$$l(1) = (0 - 0.608)^2 + (0 - 0.657)^2 + (1 - 0.455)^2 + (1 - 0.259)^2 = 1.6474$$

$$l(2) = (0 - 0.28)^2 + (0 - 0.14)^2 + (1 - 0.21)^2 + (1 - 0.37)^2 = 1.119$$

$$l(3) = (0 - 0.1)^2 + (0 - 0.2)^2 + (1 - 0.5)^2 + (1 - 0.1)^2 = 1.11$$

Step 4: The training data is the one that is nearest to output node 3. As a result, node 3 is the winner. Node 3's frequency should be changed.

Step 5: The winning unit's weights have been updated:

$$W_3new = W_3old + \alpha(x - W_3old)$$

Step 6: $Epoch1$ is finished. Reduce the proportional gain: $\alpha(t+1)$
$$\begin{bmatrix} 0.608 & 0.28 & 0.07 \\ 0.657 & 0.14 & 0.14 \\ 0.455 & 0.21 & 0.65 \\ 0.259 & 0.37 & 0.37 \end{bmatrix}$$
$=0.2\alpha(t) = 0.2(0.3) = 0.06$ and continue for successive $epochs$ unless δW_j remains stable for all training sets or the deviation is within an acceptable limit.

4. Results & discussion

TinyOS is used to generate a realistic IoT environment by establishing an IoT network with sensor nodes and random energy distribution among the nodes using the TinyOS simulator (TOSSIM). Table 5 contains a list of the simulation parameters. Various settings for the proposed network are configured utilizing network components, interfaces, event handlers, and accessible sensors. Real-time event detection is accomplished on Crossbow motes utilizing a temperature sensor in the proposed system. The system's performance is assessed using features such as threshold values, cluster formation based on residual energy in nodes, event occurrence, and traffic flow in the system for synchronous and asynchronous monitoring.

4.1. Performance evaluation results

Simulation is used to assess the success of the suggested strategy based on IoT application features such as topology, dependability, movement direction, location, total hops used, energy usage at different levels, durability, reliability, and so on. The clustering of N using GFF-NN is evaluated for varying amounts of simulated parameters.

Fig. 8 shows the FND for the number of N_i's in the network. When comparing LEACH and LEACH-C to the proposed dynamic routing protocol, it can be seen that the first node death happens earlier for LEACH and LEACH-C. This conclusion is important in comparing the proposed routing protocol to existing routing protocols since FND occurs later and hence there is an improvement in a network lifetime. The proposed protocol gained FND ranging from the 57% to 68% by varying number of nodes 50 to 100.

The LND for the number of N_i's in the network is depicted in Fig. 9. LND occurs in a greater number of rounds for the proposed routing protocol. As a result, the proposed protocol gained LND ranging from 45% to 36% by varying the number of nodes from 50 to 100. The number of packets transmitted by N_i to CH and CH to sink is shown in Fig. 10 and Fig. 11, respectively. As a result, the number of packets sent to CH and sink is increased from 30% to 40% by varying the number of nodes from 50 to 100. These data show that the proposed protocol outperforms the peer routing protocol in terms of throughput.

Because clustering in the proposed routing protocol is based on N_i characteristics, network dynamics, and application requirements, the clusters generated are very dynamic and may be employed for a variety of application requirements. Clusters and CH are chosen at random depending on the resources available in the system, network factors, and application requirements. The rotation of CH and the mobility of N_i are considered, energy consumption in N_i is minimized, and the network is load balanced. As a result, the proposed network's life duration rises 50%–60% while compared with the LEACH and LEACH-C.

In terms of network convergence time, the faster nodes converge to the most recent position of a mobile sink, the better communication takes place during the data dissemination phase. When the sink moves

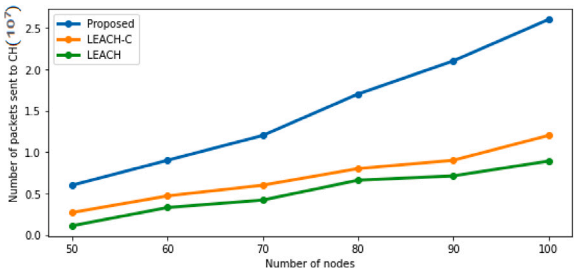


Fig. 10. Number of packets sent to CH.

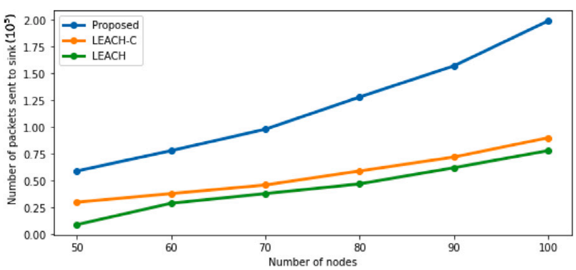


Fig. 11. Number of packets send to sink.

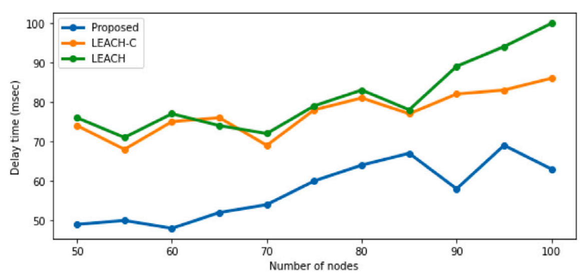


Fig. 12. Network convergence time.

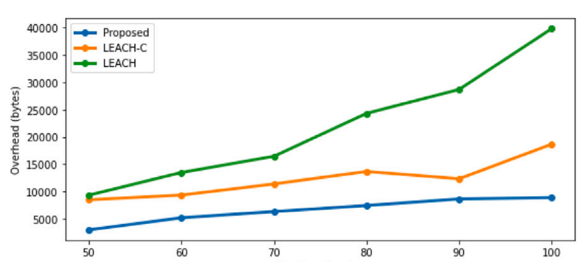


Fig. 13. Network overhead.

at a speed of 10 m/s, the convergence time of the proposed approach is effective when compared to LEACH and LEACH-C, as illustrated in Fig. 12. Using a set of neuro-fuzzy principles, the approach intelligently selects efficient nodes to collect and send the data packets to the sink node. This approach significantly minimizes network overhead and increases node convergence to the most recent position of the mobile sink.

The network overhead for proposed and peer-existing routing protocols is shown in Fig. 13. As shown in Fig. 13, the proposed scheme outperforms the LEACH-C and LEACH algorithms with reduced overhead due to dynamic clustering. Because of the uni-cast nature of the message forwarding mechanism in its standard design, the LEACH and LEACH-C protocols perform very poorly.

Fig. 14 depicts the average packet delay of the various peer-existing routing methods. The proposed protocol outperforms LEACH

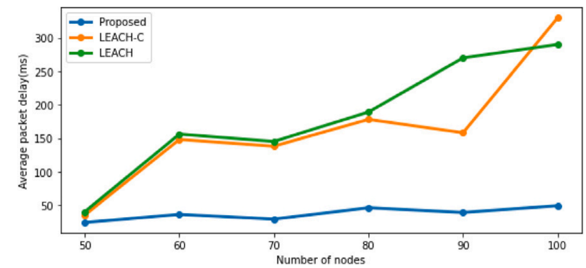


Fig. 14. Average packet delay.

and LEACH-C due to its single-hop communications while sending packets to CH and sink.

5. Conclusion & future scope

To examine cluster formation behavior depending on sensor node, network features, and green application needs, the proposed protocol utilizes neuro-fuzzy techniques for clustering and their analysis. The technique comprises using artificial intelligence to construct dynamic clustering based on sustainable IoT parameters. For clustering, the proposed protocol uses neuro-fuzzy logic for CH selection and dynamic clustering. The most significant difficulty in achieving sustainable IoT is improving the green application requirements in terms of resource management and network lifetime. So, to increase the lifetime of the network, neuro-fuzzy is applied in a dynamic clustering scheme to show the significant gain on FND, LND, number of packets sent to CH & sink, network convergence time, network overhead, and average packet delay. The future work can be dynamic node selection based on Euclidean distances such as Manhattan distance, and Chebyshev distance to improve the network lifetime of sustainable & green IoT-based smart cities.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

References

Abbasi, I. A., & Shahid Khan, A. (2018). A review of vehicle-to-vehicle communication protocols for VANETs. *Urban Environment Future Internet*, *10*(2), 14–24.

Al-Mutiri, R., Al-Rodhaan, M., & Tian, Y. (2018). Improving vehicular authentication in VANET using cryptography. *International Journal of Communication Networks & Information Security*, *10*(1), 248–255.

Chauhan, P., Chandra, S., & Maheshkar, S. (2016). Static digital signature recognition and verification using neural networks. In *1st India International conference on information processing (IICIP)* (pp. 1–6). IEEE.

Chien, H. Y. (2007). SASI: A new ultralight weight RFID authentication protocol providing strong authentication and strong integrity. *IEEE Transactions on Dependable & Secure Computing*, *4*(4), 337–340.

Chithaluru, P., Al-Turjman, F., Kumar, M., & Stephan, T. (2020a). I-AREOR: An energy-balanced clustering protocol for implementing green IoT in smart cities. *Sustainable Cities & Society*, *61*.

Chithaluru, P., Al-Turjman, F., Kumar, M., & Stephan, T. (2021a). MTCEE-LLN: Multilayer threshold cluster-based energy-efficient low-power and lossy networks for industrial internet of things. *IEEE Internet of Things Journal*, *9*(7), 4940–4948.

Chithaluru, P., Al-Turjman, F., Stephan, T., Kumar, M., & Mostarda, L. (2021b). Energy-efficient blockchain implementation for cognitive wireless communication networks (CWCNs). *Energy Reports*, *7*, 8277–8286.

Chithaluru, P., Khan, M. S., Kumar, M., & Stephan, T. (2021c). ETH-LEACH: An energy enhanced threshold routing protocol for WSNs. *International Journal of Communication Systems*, 48–81.

Chithaluru, P., Kumar, S., Singh, A., Benslimane, A., & Jangir, S. K. (2021d). An energy-efficient routing scheduling based on fuzzy ranking scheme for internet of things. *IEEE Internet of Things Journal*, *9*(10), 7251–7260.
Chithaluru, P., & Prakash, R. (2018). Simulation on SDN and NFV models through mininet. In *Innovations in Software-Defined Networking & Network Functions Virtualization- IGI Global* (pp. 149–174).
Chithaluru, P., & Prakash, R. (2020). Organization security policies and their after effects. *Information Security & Optimization*, 43–54.
Chithaluru, P., Prakash, R., & Srivastava, S. (2018). WSN structure based on SDN. In *Innovations in Software-Defined Networking & Network Functions Virtualization- IGI Global* (pp. 240–253).
Chithaluru, P., Singh, K., & Sharma, M. K. (2020b). Cryptocurrency and blockchain. *Information Security & Optimization*, 143–152.
Chithaluru, P., Stephan, T., Kumar, M., & Nayyar, A. (2022). An enhanced energy-efficient fuzzy-based cognitive radio scheme for IoT. *Neural Computing and Applications*, *34*(21), 19193–19215.
Chithaluru, P., Tiwari, R., & Kumar, K. (2019). AREOR-adaptive ranking based energy efficient opportunistic routing scheme in wireless sensor network. *Computer Networks*, *162*.
Chithaluru, P., Tiwari, R., & Kumar, K. (2021e). ARIOR: Adaptive ranking based improved opportunistic routing in wireless sensor networks. *Wireless Personal Communications*, *116*(1), 153–176.
Chithaluru, P., Tiwari, R., & Kumar, K. (2021f). Performance analysis of energy efficient opportunistic routing protocols in wireless sensor network. *International Journal of Sensors Wireless Communications & Control*, *11*(1), 24–41.
Hasrouny, H., Samhat, A. E., Bassil, C., & Laouiti, A. (2017). VANET security challenges and solutions: A survey. *Vehicular Communications*, *7*, 7–20.
He, L., & Zhu, W. T. (2012). Mitigating DoS attacks against signature-based authentication in VANETs. In *IEEE International Conference on Computer Science and Automation Engineering (CSAE),vol. 3* (pp. 261–265).
Jain, A., Singh, J., Kumar, S., Florin-Emilian, T., Traian Candin, M., & Chithaluru, P. (2022). Improved recurrent neural network schema for validating digital signatures in VANET. *Mathematics*, *10*(20), 3895.
Javed, A. R., Shahzad, F., Rehman, S. ur., Zikria, Y. B., Razzak, I., Jalil, Z., et al. (2022). Future smart cities requirements, emerging technologies, applications, challenges, and future aspects. *Cities*, *129*, Article 103794.
Jena, L., Ammoun, L., & Chithaluru, P. (2022). Supervised intelligent clinical approach for breast cancer tumor categorization. In *Augmented intelligence in healthcare: A pragmatic and integrated analysis* (pp. 15–40). Springer, Singapore.
Joshi, D., Chithaluru, P., Singh, A., Yadav, A., Elkamchouchi, D. H., Breñosa, J., et al. (2022a). An optimized open pit mine application for limestone quarry production scheduling to maximize net present value. *Mathematics*, *10*(21), 4140.
Joshi, D., Chithaluru, P., Singh, A., Yadav, A., Elkamchouchi, D. H., Pérez-Oleaga, C. M., et al. (2022). A novel large-scale stochastic pushback design merged with a minimum cut algorithm for open pit mine production scheduling. *Systems*, *10*(5), 159.
Manman, L., Xin, Q., Goswami, P., Mukherjee, A., & Yang, L. (2020). Energy-efficient dynamic clustering for IoT applications: A neural network approach. In *2020 IEEE Eighth International Conference on Communications and Networking (ComNet)* (pp. 1–7). IEEE.
Mukherjee, A., Goswami, P., Yang, L., Yan, Z., & Daneshmand, M. (2020). Dynamic clustering method based on power demand and information volume for intelligent and green IoT. *Computer Communications*, *152*, 119–125.
Musaddiq, A., Zikria, Y. B., Hahm, O., Yu, H., Bashir, A. K., & Kim, S. W. (2018). A survey on resource management in IoT operating systems. *IEEE Access*, *6*, 8459–8482.
Ramakuri, S. K., Chithaluru, P., & Kumar, S. (2019). Eyeblink robot control using brain-computer interface for healthcare applications. *International Journal of Mobile Devices, Wearable Technology & Flexible Electronics (IJMDWTFE)*, *10*(2), 38–50.
Saleem, A., Afzal, M. K., Ateeq, M., Kim, S. W., & Zikria, Y. B. (2020). Intelligent learning automata-based objective function in RPL for IoT. *Sustainable Cities and Society*, *59*, Article 102234.
Tanwar, R., Balamurugan, S., Saini, R. K., Bharti, V., & Chithaluru (2022). *Advanced Healthcare Systems: Empowering Physicians with IoT-Enabled Technologies*. John Wiley & Sons.
Wang, X., Li, S., Zhao, S., Xia, Z., Bai, L., et al. (2017). A vehicular ad hoc network privacy protection scheme without a trusted third party. *International Journal of Distributed Sensor Networks*, *13*(12), 155–175.
Zikria, Y. B., Kim, S. W., Hahm, O., Afzal, M. K., & Aalsalem, M. Y. (2019). Internet of Things (IoT) operating systems management: Opportunities, challenges, and solution. *Sensors*, *19*(8), 1793.
Zikria, Y. B., Yu, H., Afzal, M. K., Rehmani, M. H., & Hahm, O. (2018). Internet of things (IoT): Operating system applications and protocols design, and validation techniques. *Future Generation Computer Systems*, *88*, 699–706.

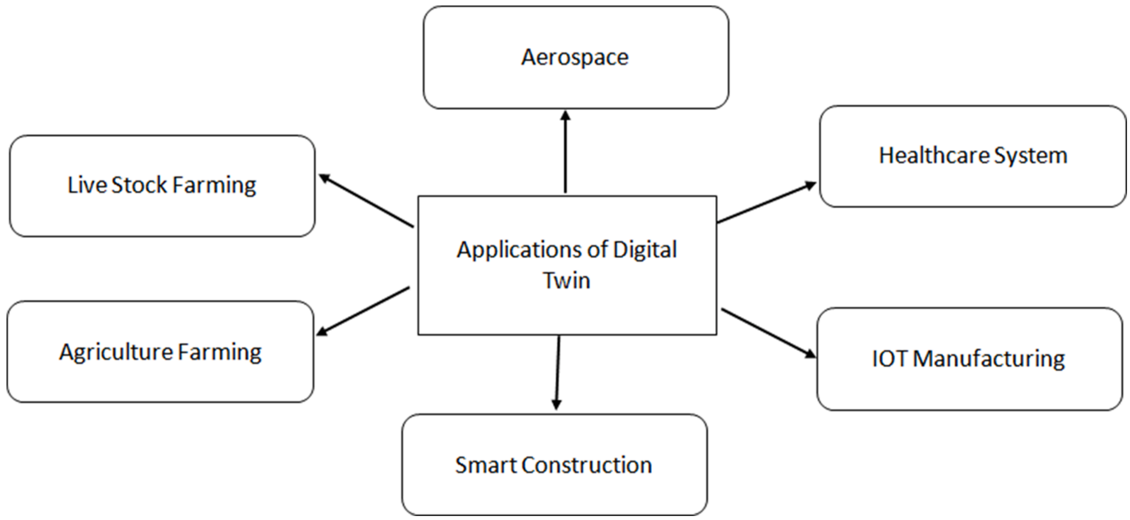


Fig. 1. Application of Digital Twin in An Advanced Environment.

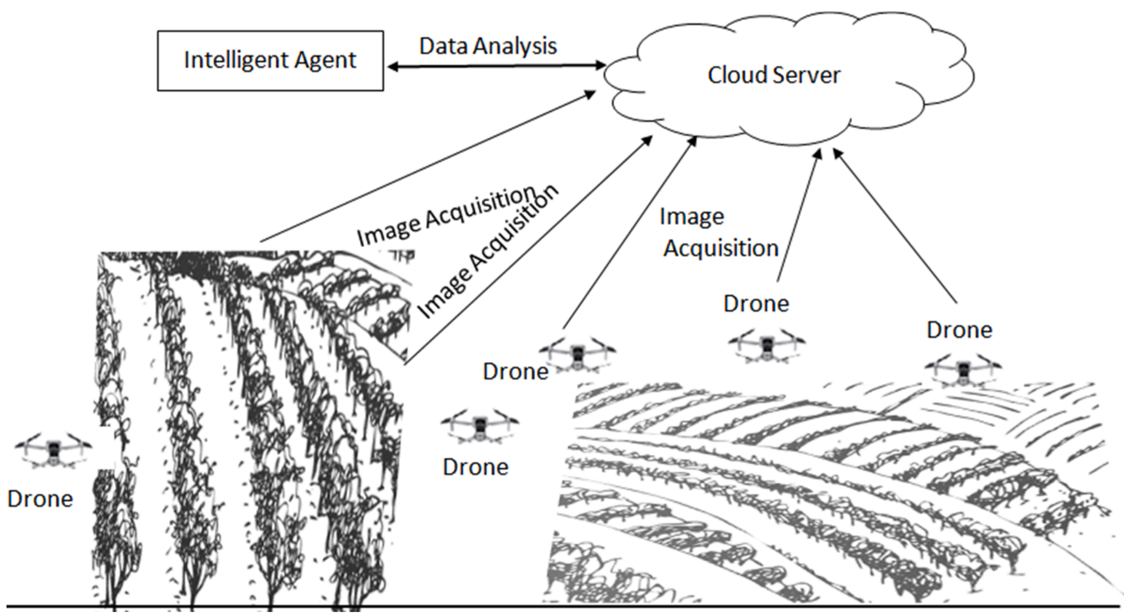


Fig. 2. Smart Agriculture Farming Using Digital Twins in Intelligent Agent Module.

product data analysis and support remote service [5]. The actual education training and monitoring through virtual environment especially in the covid-19 pandemic. The somato sensory method is used to recognize exercise practice efficiency. Kinect sensors are used to know the bone movement details, which the card reader will use to obtain personal information age, gender and weight. We use motion Recognization using the gain algorithm. The data captured is stored on the cloud server. All the information exchange is done through digital twins. This approach is used to monitor students [6].

2. Related work

Digital Twin for agriculture will help for low price cropping help to produce high productivity using the technology of AI and the Internet of things for intelligent monitoring. The wireless Sensor framework is designed and connected to a network. Sensor data is stored on cloud devices using AI & ML methods to detect crop disease by observing the images of the leaf taken by drone cameras. Monitoring the field through object detectors and survivance cameras. It also observes any plant disease, wild plant growing in the crop, or plant nutrition deficiency are some of the digital twins and sample illustration of DT is shown in Fig. 2 [7].

Digital livestock Farming uses advanced computer applications with IoT concepts to monitor by creating an artificial virtual world and operating remotely. Livestock farming help in farming practice for wide varieties of animals and examine animal feature analysis like eye movement, facial expression, ear analysis, predict the heat cycles and observe improper behavior of animals. The major Utility of Digital twins is to know the animal's emotions, maintain the climate constant using heaters and air conditions, Movement of animals using GPS location monitoring, and analyze the animal's growth and proposed model illustration is shown in Fig. 3 [8].

The DT concept application in industries, especially manufacturing, makes a finished product from the raw materials in a simulated model. The entire life cycle phases of manufacturing products are monitored, assisted, and checked. The purpose of the DT is to design the product using augmented reality to predict the actual maintenance, and data analytics are used for zero defect manufacturing. The author discussed

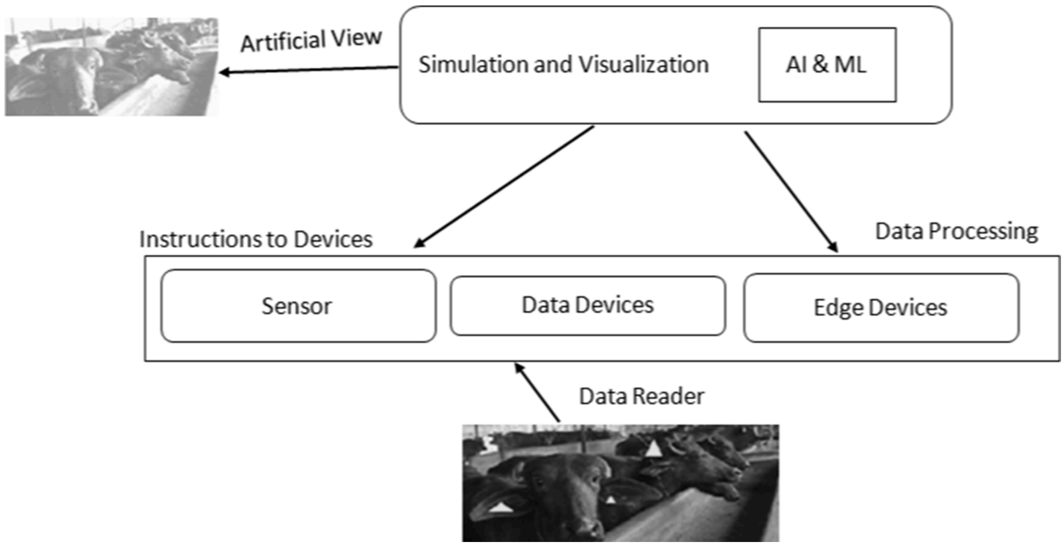


Fig. 3. Smart Livestock Farming Using Digital Twins Simulator.

the digital twins helped in different phases of manufacturing of Finite Element Modeling in chain production. The sample results are projected in the graph. The time varies as per the temperature used for the material for shrinkage [9].

Autonomous Driving using digital twins helps provide security, assess, and validate road safety in automated driving. Use cases are designed for the identification of objects, analysis of data, security analysis, and safety measures placed on the real-world environment. The real threat assessment is asses on the virtual world using risk prediction using machine learning concepts. The SCARI Approach will observe the data, orient the analysis, decide (plan maker and decision), act by the event, and update the knowledge base. The metrics used are a digital signature, Round trip time, and packet send and receive rate. The IoT4CPS project helps people to assess the risk [10].

DT in health care systems named as DTH data has an actual entity, virtual thing data, service information and their fusion data, i.e., dynamic, bi-directional links among the actual entity and its equivalent twining take place in a digital world [11]. The three DTH data components include; (1) a virtualization core that can solute higher difficult math problems demonstrating actual models has dynamic conditions (2) innovative methods that activate data, content creation and information-driven modeling; (3) devices to accepted models to explore in environments online [12]. Personalized medicine or healthcare starts from the hypothesis that modified math models of Subjects, powered with large profile information, will help choose apt and more effective medical intrusions [13]. DT technology can initiate a significant transformation of traditional data management to automated medical records (individual/s), and their collections (overall population) are available for accurate medicine [14].

DTH has complete information about the human body or certain organ systems or organ functions, such as the digestive system/function, liver, or body components such as cellular, and sub cellular levels involving – organelle or sub-organelle and molecular levels. This is further utilized to design for body organ disease, disorders and some diseases, e.g., arthritis, nervous disorders like Multiple sclerosis etc., liver with non-alcoholic liver disease, - Coronary artery inflation problem, Pericardial issue, Cardiomyopathy, Congenital heart disease etc. for infections by different microbes like bacteria, fungi, protozoan's, viruses etc., interacting with the human body [15].

DT can equally be beneficial to both healthy or diseased conditions, for instance, a diseased cell, such as a cancerous cell and its types (carcinomas, sarcomas, leukaemia sand lymphomas), and organs affected (kidneys, eyes, heart, blood vessels, nerves) in type 2 Sugar, a fatty liver condition, a syndrome affecting the whole body/multiple parts or autoimmune diseases [16]. IoT has offered widespread assistance in various disciplines in industry, similarly in the healthcare sector, right from the collection of real-time data from linked clinical, health, environmental sensors, etc., and devices to simplify the communications between equipment/machines and humans. This helps bring important data available through electronic medical records, diagnostic procedures, remote monitoring, and patient reports [16].

The technologies accompanying artificial intelligence, including machine learning, are promising, as they offer cutting-edge data analytics and cloud computing. With on-demand networked computational resources, it also becomes a crucial tool for processing large quantities of data and discoveries, which are helpful in real-time [17]. The blend of these tools/technologies has an incredible collaborative effect by generating timely and valuable insights for medical professionals and individual patients; likewise, improving them will be more advised in taking proactive decisions. These also serve as the endurance for shifting healthcare towards more precision, personalized treatment management and preventive care [18].

The significance of research carried out in various center efforts in starting precision cardiology using cardiac digital twins (CDT) [19]. The cardiovascular system being the most critical system among various systems of the human body, the CDT model is anticipated to maximize the interaction amongst the anatomical, mechanical, and functional entities of the cardiovascular system and advanced analytical models created around related data, with an interpretation of descriptions to predictions of conditions [20]. The 'Living Heart', published software, can transform an individual's two-dimensional (2D) scan into a full-dimensional heart model, delivering users resources to manipulate the virtual heart model. A current study recommended a plan for the generation of automatic CDTs that highlighted the reliability of the simulated model and constructed computational proficiency [21].

The increased range of applications in the biopharmaceutical industry - for drug discovery and development involving biosimilar, hormones, anti-cancer, or generic drugs. DT of the liver involves integrating knowledge and understanding of the various liver functions, diseases, and effects of various drugs, using a framework of ordinary differential mathematical equations. The system coupling liver DT in conjunction with experimental measurements (values) has been demonstrated to understand drug-induced liver damage or toxicity. Similarly, studies on other diseases, organ dysfunctions, or cancers can be investigated, and analyzed, or drug management or discovery can be created using these applications [22–24].

With the advancement in various tools/technologies, such as big data, cloud computing, and the IoT, the DT has been employed in industry to transform concepts into practices by precision simulation technology [25]. Simulation also has an important role in healthcare, especially in research conducted in medical path development, allocation of medical resources, prediction of medical activity, etc., [26]. As a result, utilizing DT in the healthcare industry will be an innovative and cost-effective way to offer more precise and quick services for elderly/physically challenged healthcare [27]. IoT influences healthcare services effectively, in diverse valuable ways – by supporting hospital/s organization and management, for medical pathway planning, in medical or quality resource provision, for better diagnostics or treatment and care, and by predicting patient outcomes and disease progression categorizing personalized therapies [28,29].

Healthcare system or assist resources comprise two components centered on hardware and software [30]. The hardware resources are primarily represented by professional healthcare equipment such as computed tomography (CT), sonogram/ ultrasound machines, gamma cameras, magnetic resonance imaging (MRI), and various auxiliary equipment involved in physiotherapy [31]. The professional software resources are matched with different healthcare equipment. A health information system (HIS) involves sharing protected health information (PHI) between organizations and providers hassle-free and for efficient patient management. The clinical information system (CIS) is at a single location for the entire healthcare, providing patient data and management. An electronic health record system (EHR) is a real-time, patient-centered record helping to understand patient medical history [32,33].

Healthcare competencies are formed with diagnostic facilities, collaboration capacity, rehabilitation care ability, expert knowledge, and intellectual reserves, e.g., clinical trials facility, professional healthcare records and advanced healthcare equipment [34]. The information includes actual activity monitors and smart wearable's, bracelets, step counters or sphygmomanometers, and other elderly or specially challenged [35]. The smart healthcare public service sector results in the design of applications to be user-friendly for both medical professionals and patients for maximum advantage. The levels of smart healthcare are divided into three categories based on the requirements – The Hospital and Health Care Centers, The Medical Research Centers and Clinical Trials department Local Health Department and Health Minister & personal people [36].

Digital Twins is a transformation of technology in a new dimension and enhancing the research opportunities to solve the complex problems in various industries. DT is an integration of multiple technologies, coupling with each other to perform extensive task. For every country agriculture cultivation are the roots for country wealth, health and growth. Many applications are being developed using digital twins in agriculture. The author has discussed the opportunities and feasibilities in agriculture using digital twins serving the services like dynamic monitoring system, energy consumption, failures of machine, virtual operations, predict the behavior of mammals, modification of parameters [37].

Digital Twin is creating a new era in technology convergence for industrial applications. DT is addressing the complex problems by using the latest production equipment, electronic devices and sensors. The latest technology like cloud, edge, Fog computing, IOT devices and another major paradigm are merged with digital Twins in design development, processing and services. The vision of digital twins is to provide connected architecture from inception to product service in one complete life cycle, it consists of data processing, modeling and services [38].

Digital twin healthcare system for 4.0 industry has various services like Digital Assister, Personal Health Diet Monitoring System, AI based Rehabilitation System, Drug Development, Trauma Management, Treatment Assister and Diagnosis Assister. Digital Twin Architecture for Health Monitoring System Architecture has been designed with convergence of technologies like Machine Learning, Big Data Storage and Data Sources are communicating with external entities of the real world like Healthcare Community, medical services department like ambulance, Chatbot, Robot care, Robots in labs. Technical Support like Medical Applications, Smart Gadgets, Tele-Health Systems [39].

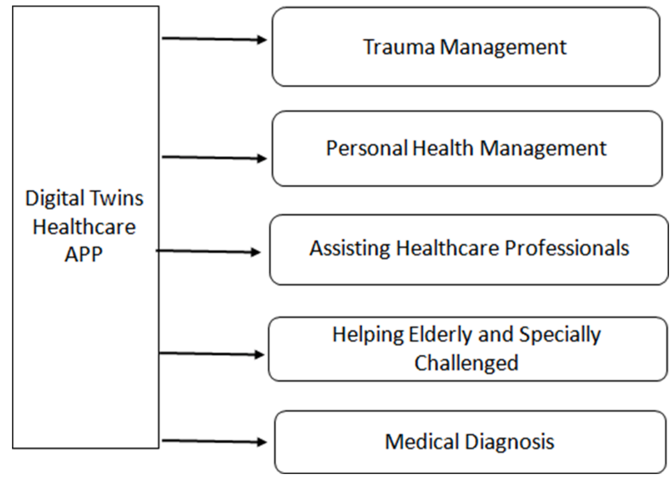


Fig. 4. HealthCare Applications using Digital Twins.

3. Digital twins technology in healthcare system

3.1. Digital twins aiding diagnosis and treatment

The tools such as artificial intelligence (AI), automated machines, surgical robot, and diverse authenticity, the diagnosis and treatment has become extra intelligent in the clinical decision in conditions. The diagnosis of various diseases or disorders like hepatitis, lung/skin cancer, Chronic Obstructive Pulmonary Disease (COPD), Multiple Sclerosis etc., with high accuracy than experienced physicians [40,41]. IBM's Watson support system for clinical decisions is the best outstanding invention so far. The intelligent cognitive system's clinical data has effectively analyzed cancer and diabetic diagnosis and some of the applications are shown in Fig. 4 [42,43,36].

3.2. Digital twin assisting personal health management

Chronic diseases/illnesses are incurable and costly; as a result, the health management of the disease is very important. Smart healthcare is a new health management model providing extensive attention to patient self-management emphasizes real-time self-monitoring of patients, instantaneous feedback on healthcare data, and well-timed interference of therapeutic performance [44,45].

The rise of implantable prosthetic devices or smart wearable gadgets, smart appliances in homes, and smart tools/ programs in health information connected through IoT. In divergence, third-generation devices are advanced with the usage of electronic components like biosensors, IoT devices and Integrated Microprocessors to monitor applications of patient data regularly, and Artificial Intelligence is used. The Utility of the power is very low compared to other alternate electronic devices. It is easy to use and allows collaboration with other applications, DB. These monitoring devices regularly will help control the disease attacks and prognosis status with high accuracy [49].

The detection of blood glucose monitors the changes in glucose levels and helps in diet planning, actual activities, lifestyle changes, and medication time. An electrocardiogram (ECG) detects blood pressure, a signal of pulsation and pressure with sensors. Checking the saturation of oxygen and the body temperature became a basic need in the current pandemic, not just in intensive care units; therefore, it's important in current smart healthcare [50]. The unseen or most avoided health issue

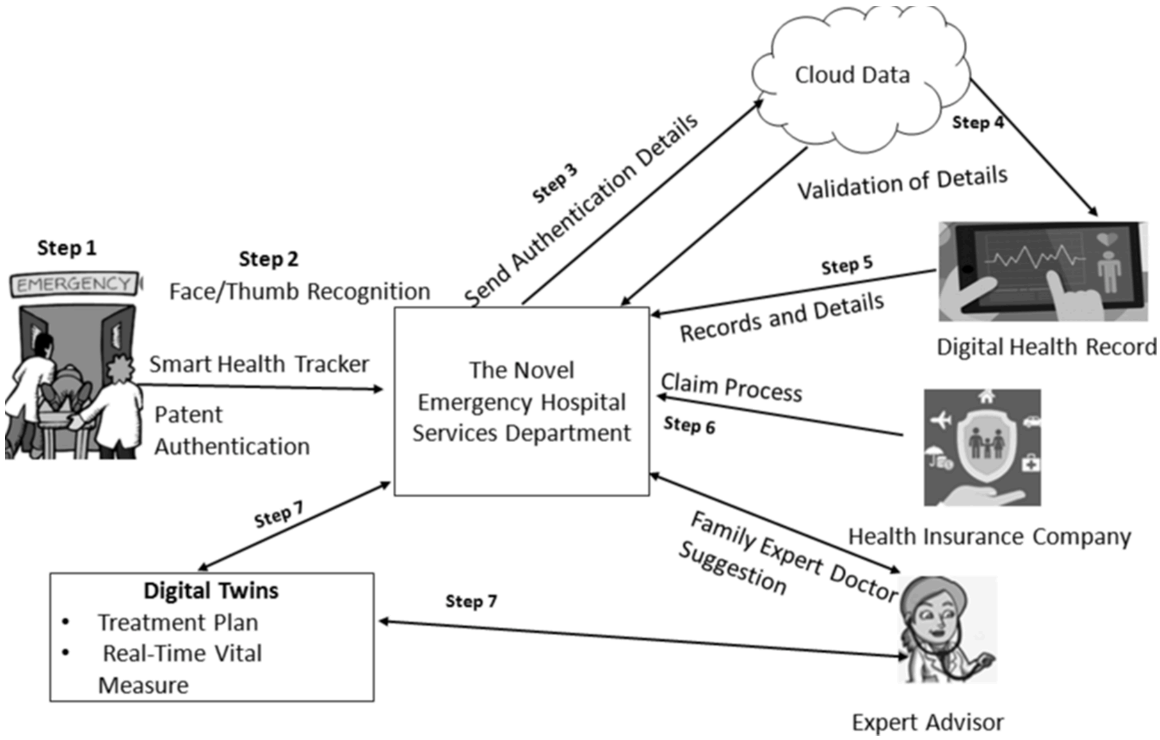


Fig. 5. Novel Approach for Emergency Hospital Service (E-H-S) using Digital Twins.

in the current scenario is psychosomatic conditions, where smart healthcare management can aid in self-manage for individual stress detection and alleviation system. Wherein continuous monitoring automatically helps in reducing stress by alerting the person with warning signs, integrating biosensors into their phones, watches/fit bands, providing an innovative model of monitoring their body or environment and improving portability, with high performance for users more easily [51,52].

3.3. Digital twins in helping elderly and specially challenged

Smart residential homes support the elderly and the specially challenged or differently able individuals, with infrastructure integrated with sensors and actuators to monitor the homes, actual signs and environment in the absence of the caretaker or patient attendee. The automatic monitoring system would help perform operations that improve living experiences and homes or healthcare. Hence, we can confine to giving services while acquiring the data processes by reducing the dependency on healthcare professionals and improving their quality of life [53,54]. It is promising to put all the collected health data from multiple portable devices into a clinical assistance system to generate a classified health decision aid approach that can fully access the collected data for accurate diagnosis [55].

3.4. Digital twins in trauma management

DT starts even before the patient arrives in the hospital, which is very important for the case study. Therefore, the trauma team will be trained to alert the incoming patient and start collecting and receiving information right away from the accident site/injury site. The internal state or vitals of the patient changes rapidly while transferring to the emergency department. The trauma management role involves the patient information from all the connected devices (i.e., the vital signs monitor), trauma leader, shock-room, room of the emergency department, other tools, and equipment (e.g., rapid diagnostics machinery, displays for real-time information of the ongoing trauma etc.) [56].

3.5. Digital twin in assisting healthcare professionals DT

It can be very helpful in assisting clinical decision-making, predicting possible risks associated with patients' underlying conditions, and recommending in advance. The latest proposal idea will help all the stack holders of the healthcare system like scientific health researchers, Subjects, and doctors will help to remove the objections in the collaboration process. These permit patients to approach telemedicine advice and facilities effortlessly; doctors can also vigorously observe patient status remotely. The medics are well equipped with experts and researchers and by collaborative approach for better outputs [57]. The digital networking architecture can improve the overall system by reducing medical mistakes and issues in medical treatment, improving the prompt time of medical emergency service, and providing healthcare services considering the common person's pockets [58].

Digital (smart) healthcare management comprises of following important elements, such as Information Communication Technology (ICT) to maintain accurate information provided by the hospitals, Family Members, and Regional Health Departments, particularly those established by IoT optimizations and automatic procedures, to enhance current subject care practices by introducing novel features [59]. The smart hospitals integrated with DT may include operations to take care of medical carers, non-medical staff, patients, and administrators. The challenges faced by the users need to be considered by hospital management and decisions made based on them. ICT platform integrated with multiple digital systems built by IoT that connects digital devices, staff management tracking instruments, intelligent buildings, personnel, and biological specimens [60].

3.6. Digital twins in other managements

The digital healthcare system has been used in the pharmaceutical industry and biologics for inventory management, production, anticounterfeiting circulation, risk management, and security. DT in scientific innovation and development for further precision & accessibility, including clinical trials, target screening, drug discovery, and finding

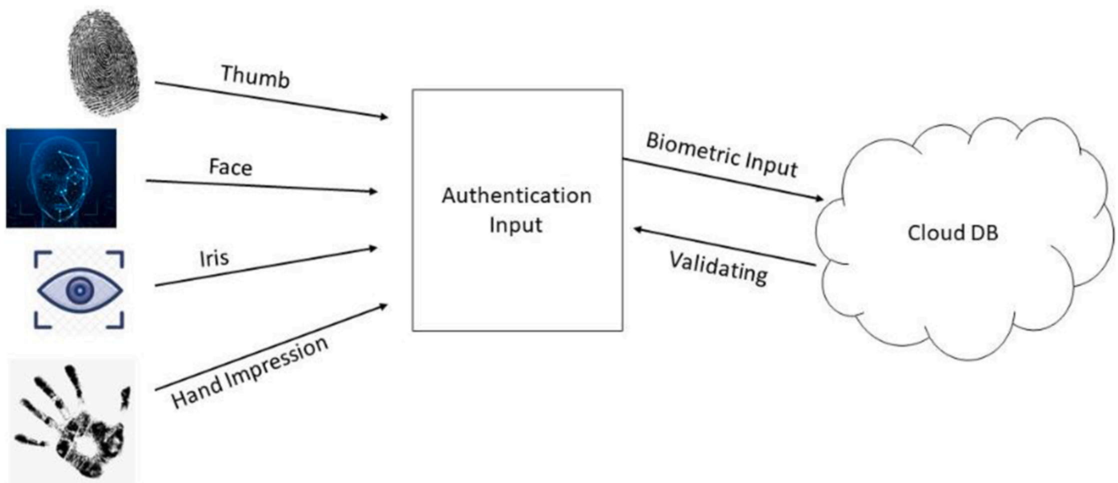


Fig. 6. Biometric Authentications on Cloud Environment.

effective target points. A few examples include identifying targets for the ribonucleic acid-RNA imperative of Proteins in amyotrophic lateral sclerosis using the Watson system for genomics studies on tumours. The DT system can also gather updated information from worldwide research in dynamics and will help optimize or correct the screening process at any stage [59,61].

4. Novel proposed approach for emergency hospital services

The proposed novel idea has been designed for the patients who joined the hospital for emergency services, without carrying much information about the patient health track record, illness details or maybe accident subject, a person can be anonym etc. the motivation for my proposal of emergency treatment is always key for the risk patient time, treatment plan, basic information, and family doctor opinion makes the situation more controllable and manageable to treat patients at risk hours. The blood group and basic diabetic history, early disease and issues need to be digitally recorded by every person and allow access at emergency. The digital twins' technology can be used to know the basic health digital record of even unknown patients. The primary study helps the doctors at risk hour treatment. The technology needs to be maintained by people for personal health care details and maintain the track record in their digital storage and in Fig. 5 illustrated the flow of Emergency proposed model.

Digital Health Record (DHR): The patient's Digital Record needs to be maintained on the cloud with face or thumb Recognition to access the data over the cloud. The records contain patient age, gender, blood group, minor and major health issues, operations, allergies, diagnosed reports, master health check reports, insurance claim details etc. Smart devices Health Tracker (SDHT): The wearable devices are smart enough to track the patient's details like heart rate, body temperature sensor, no of steps walked and actual activity to know the lifestyle of living. The latest 48 hrs of patient data can be seen by doctors, as smart wearable devices store data on the cloud for later retrieval. Expert Advisor (EA): The family doctor knows the in and out of the patient details, treatment undergoing and responsiveness of the patients. The advice from the doctor will surely help the emergency service team. The personal expert adviser's doctor opinion can be added to the treatment plan. The information to a family member and insurance company for fast processing.

Digital Twins: Digital Twins are used to projecting the medical report virtually to an expert adviser, Emergency Hospital Service. Access to Smart Health Tracker information is updated on the cloud and viewed by both peoples. The treatment plan can be operated using Digital Twin data sharing. The vital Measurements can be shared in a real-time environment.

Proposed Algorithm

Algorithm for Emergency Hospital Service
Step 1: **Start**
Step 2: Patient in Emergency Room (High Risk Patients).
Step 3: **Input**: Biometric/Face recognition capture [62].
Step 4: **Authentication:** Cloud Authentications for input.
Step 5: **Access**: Electronic Health Records able to access.
Step 6: **Communicate:** Cloud Communicate with E-H-S Dept, Insurance Company, Family, Family Doctors, Expert Advice for treatment plan.
Step 7: Digital Twins helps in co-participate and interlink with each other.
Step 8: **Emergency Health Service** in coordination with different sectors in virtual reality service to treat an annoy patient.

5. Methodology

5.1. *Biometric authentication on cloud computing environment*

The patient's data is most important and private in the medical industry. There is a need to secure the patients' information to the outside world and prevent misuse. The Digital Personal Health Record is maintained on the cloud for easy storage, access by any person, access at any time and from any location. The secured cloud data can't be accessed until authentication credentials are known to us. The emergency patient may not be in a position to reveal the credentials. The solution is to use biometric data for the person's authentication. The different types of biometric Recognition are Iris Recognition, Finger Geometry Recognition, Face Recognition, Typing Recognition, Hand Geometry Recognition and Voice. Speaking identification is used to authenticate the actual presence of a person. In the proposed approach, Finger Recognition and Face Recognition methods authenticate a patient in the emergency room and in Fig. 6 possible biometric authentications services in emergency shown [63].

The processes of Biometric Authentication need two-stage processes. In Stage I: (Enrollment Stage) of the person's Sensors are used to record the Thumb or Face Recognition, The Quality Assessment used for checker, Pre-Processes the data, Feature Extraction is performed, and Biometric Database is created. In Stage II: (Authentication Stage), Patient Face and Thumb are recognized, quality checker for proper assessment, Preprocessing the image, feature extraction to identify the RoI, Pattern Matching from the Biometric Database [64].

5.2. Digital twins technology in health care

Digital Twins is a collaboration of various industries, technologies and utilities. The digital twins evolved and created a virtual world. The remote access to perform operations. The digital Twin has applications development in various sectors like aerospace, manufacturing, Health-Care. The Digital Twins are used in the Health care sector to diagnose, Drug Development, Virtual Surgery, Monitoring, Medical Devices and Regulatory systems. The proposed novel method uses digital twins to diagnose the disease and treatment Decision. In the case of cardiovascular health problems, we can use heart image and computed fluid dynamics metrics to evaluate the disease.

The E-H-S in our model will use digital Twin's application based on the treatment needed for emergency patients. The Surgery Simulation will minimize the operation risk to a doctor. Like Tumor region selection can be done either by a doctor or the robotic machine, the accuracy is greater in the robotic operation than the doctor. It uses AI technology to select the tumor position, Dim and orientation. The selected knife is positioned by an angle, shape, and Magnitude system. The regulatory decision making is always easy, effective and safe, are used using wearable smart devices and monitored on the applications and notifications are created to doctor and patient, if beyond the threshold values [19,65,66].

Table 1
List of medical applications using digital twins.

List of Medical Applications using Digital Twins			
Author	**Utility Of Application**	**Methods Applied**	**Remarks**
Corral-et.al [19]	Precision cardiology Diagnosis Process and Treatment Plan	- Exploring input data - Predictive models are built using maximize the value of the data. - Inductive Reasoning &Deductive Reasoning. -Evidence Reasoning	Observing the Data Sources like Medical Images, Lab Reports, Personal Subject observations, and genotyping for diagnosis.
Schwartz [46]	E- Self Monitoring	Mobile Application Wearable Devices Medical Devices Habitat Monitoring	Used to feed with real-time data onto DT in the cloud to understand disease progression and continuous patient data collection.
Morrison [47]	Regulatory Science	- Regulatory decision making -Medical devices design and optimization. -Surgery simulation: surgery risk assessment	DT helps improve a device's performance by running hundreds of simulations with different conditions and different patients. Further, with the emergence of 3D printing technology.
Pappalardo [48]	Silicon clinical trials in drug development and dosage optimization	In silico clinical trial will focus on processes that take years to be observed *in vivo* or assess the risk of rare cases	To create a digital cohort of real patients with different phenotypes, which share symptoms, and test new potential drugs to predict with one has more possibilities to success and the optimal dosage.

5.3. Digital twins used by expert adviser

Expert Advisor can be through Expert System or Communicate to Medical Expert through Virtual World, and Advance Surgery is planned and executed through analysis and virtual systems operated and had very successful stories. The expert systems provide suggestions and recommendations for critical cases [67–69]s. The Expert system designed with machine learning algorithms and training with huge datasets and intelligent decision-maker algorithms will help take advice from the expert system [70] as shown in Fig. 7. The sample Expert system was designed to detect breast cancer using Wisconsin Dataset; the algorithms used are Association rule and Neural Network Concepts. The success ratio is 95.6%. This system has a faster detection rate than any other disease [71].

5.4. Digital twins using block chain for communications

Digital Twins use blocks chain concepts, especially at the time of emergency. The hospital people will not have time to communicate information about the pre-requirements to communicate with various departments personally. We use technology to update the patient's Details, Treatment plans, Insurance Claims, Health Reports to Personal Health experts, Insurance Claims Companies, Family Members etc. The digital twins use the proposed model to update all the partners' system changes and information updates using block chain [72].

6. Advantages of proposed model and limitations

Digital Twin concepts are implemented in the latest applications to test the model with an artificial Envioursment, continue monitoring, and simulate the design models. Digital Twins in health care has wide applications like Virtual Treatment, Electronic Health Record, Staff Training, Drug Detection, and Vital Monitoring are general applications for digital Twins. The proposed model has the intense advantage of using digital twins from anonymous patients till the person is discharged from the hospital. Identification of the patient and basic health records is very important in ER room. The proposed model helps you to identify the person with biomarkers and communicate with Electronic Health Records. The proposed model helps to identify a person's health record details and communicates with the personal doctor, expert advice, Family, and the health insurance department. Hospital Services may require at any time and location. Digital Twins can help hospitals to provide advanced treatment.

7. Results & discussion

The Novel Proposed Approach for emergency service in hospital has motivated me to help the Anonym person fall suddenly sick, had an accident, heart attack etc. It is very difficult for the E-H-S department to deal with anonymous patients and unknown their previous history. The Biometric markers authenticate to access the person's Digital Health Record in emergency cases. The empirical results prove and concern the proposed model to be success with higher percentage. The digital health record and historical study or content management system has 413 records over one year duration and this network data has successfully utilized the Electronic Health Record data collected in Emergency Department, Intensive Care Unit and Floor Departments. From 65.9% of morning data and 34.1% of night data is collected [39].

The existing results obtained are the rate of access DHR is only 19.8 s and verified for 200 K people and added advantage is before ambulance reaches the hospital pre-history study will be done by E-H-S department [62]. The Images are sent to the Cloud authentication using the Minute Map algorithm, and data shared is encrypted using the SHA algorithm. It is proven that it is possible to provide data on multiple clouds with efficient and secure Data [73]. The emergency healthcare application needs to frame an expert system using different algorithms like

Table 2
Efficiency of the proposed Model in comparison to Emergency Requirements.

Emergency Patient Requirement	Proposed Model Satisfying
Emergency Patient Requires Advance Ambulance First–Aid.	Every ambulance is provided with IoT Medical devices connected to E-H-S dept.
Biometric Authentication Detection System.	Every Ambulance or Emergency Room has a Biometric detection device to authenticate and access the medical Track records.
Communication to Family, Personal Expert Doctor & Insurance Company.	Biometric Authentication will notify the nominators to know the information.
Financial Assistance	The insurance Company will Sanction the amount based on emergency.
Basic Investigations like Blood group	Extracted from the Digital Health Record
Personal Expert Advice	Family Doctors can help in deciding on the treatment plan.

Table 3
Limitations for proposed model.

LIMITATIONS
Patients need to Maintain Digital Health records (DHR)
Permission to Access DHR based on Biometric Detection
The hospital needs to be authorized to access cloud DHR.
Insurance Companies and hospitals should have Association for sanction /claim transactions.
Family Doctors or Expert Adviser availability check in emergency for Treatment Plan.

rule-based analysis or other AI and ML methods to take an expert system decision, Expert adviser, and executives. With taking a dynamic and real-time decision [74]. The advantages and considerations to execute the proposed model are discussed in Tables 2 and 3.

The Digital Health Record conducted a survey for 170 members staff training, in Fig. 8 the statistical analysis result is represented in graph, The members responded to strongly agree is 98.8%, the quick and easy access for digital health record is 98.2% and barrier to implement health record in recording digital health record is 96.4% [75]. Internet failure are the common reasons not able to capture the data. The emergency room services play an important role to save the life of the patient in critical situation. The level of nurse knowledge and related patient information and fast responses was analyzed in Porsea Regional Hospital in 2022 and graphical representation is show in Fig. 9. The good knowledge with fast response time is 64.7%, good knowledge with slow response time is 35.3% and lack of knowledge with slow responses time is 9.7% [76].

The Biometric thumb registration are created by marking the thumb image into blocks, the core points are detected and imaginary structure are created. The new test data will perform the feature extraction with similarity comparison over inter and intra instances in a dataset. The success rate is represented in Fig. 10 for intra and inter instances in database [77]. Expert System will take decisions based on the fuzzy logic rules, correctly classify percentage is above 80%. The success rate is represented in Fig. 11 for General Patients, Normal and Abnormal Patients.

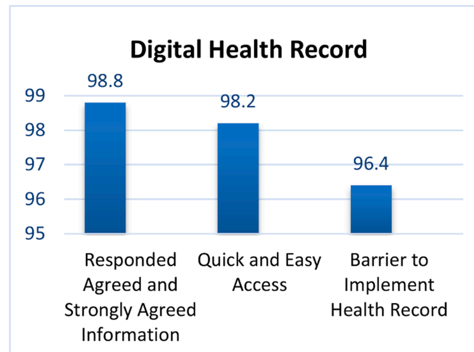


Fig. 8. The Statistical Analysis for Staff to Agree.

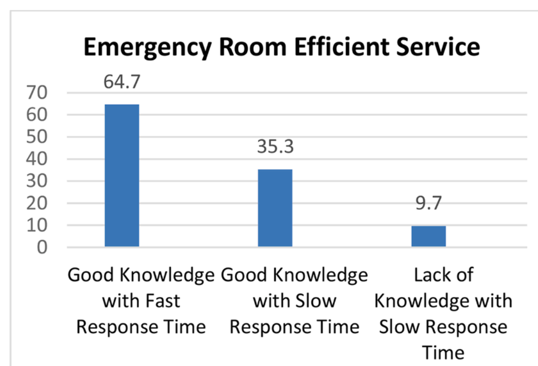


Fig. 9. The Knowledge Analysis Emergency Room Service to Record Digital Health Record.

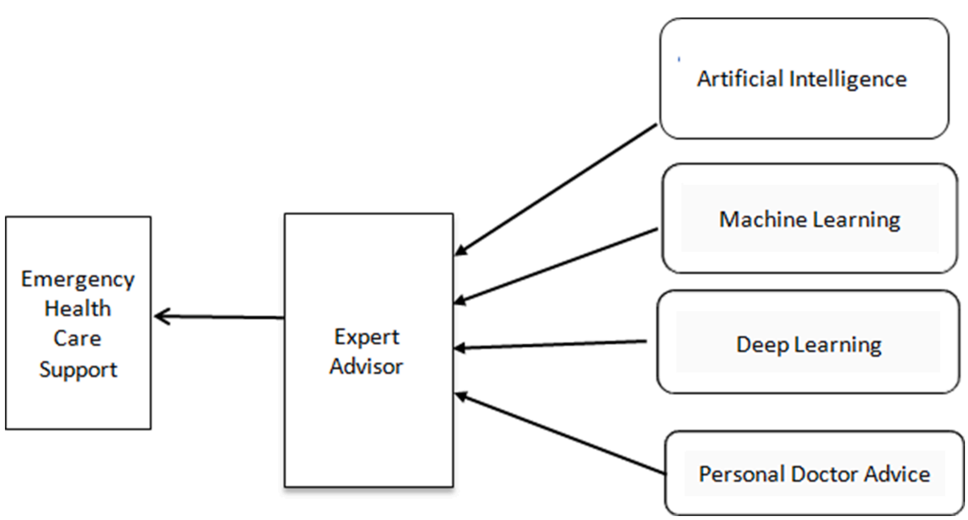


Fig. 7. Expert Advice workflow and Technical Exercising in decision making.

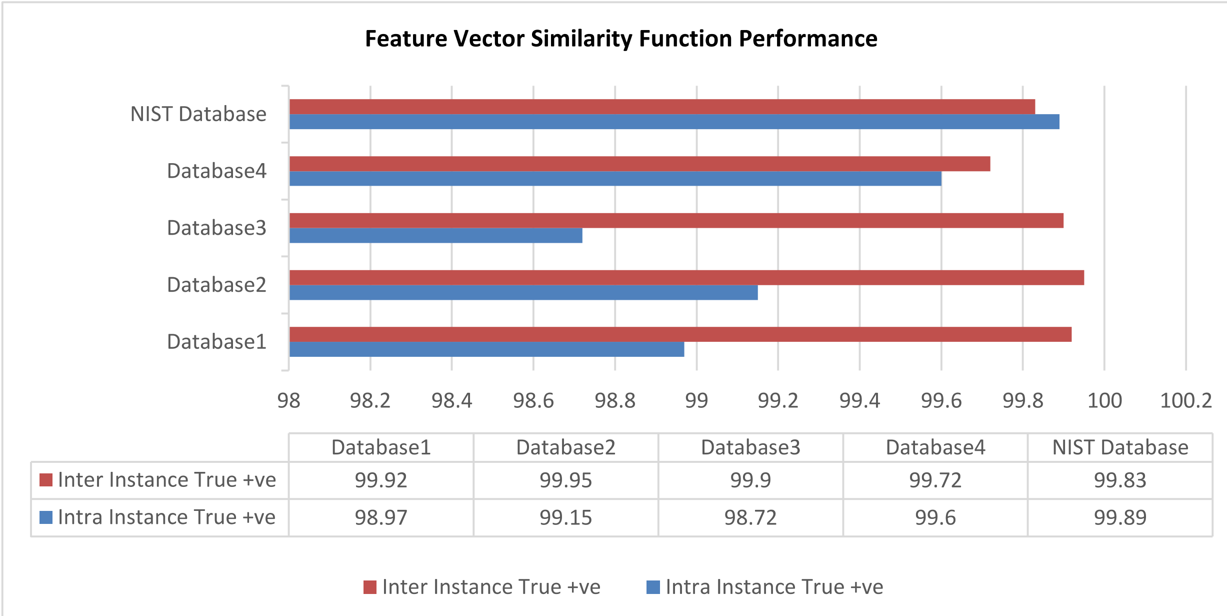


Fig. 10. Inter and Intra Instances True Positive for Similarity Feature Extraction.

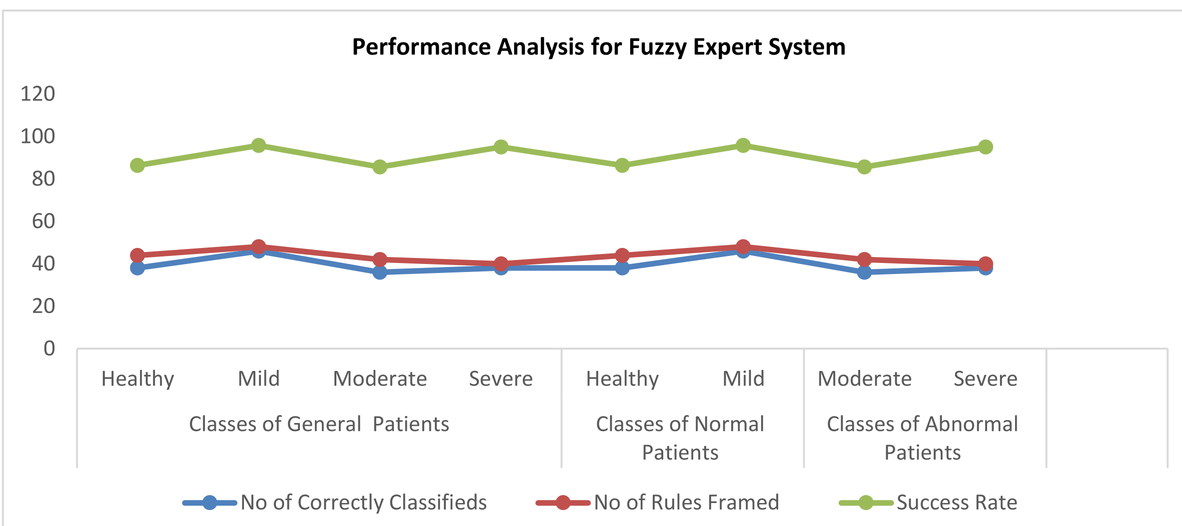


Fig. 11. Expert System using Fuzzy Logic Based Rule.

8. Conclusion & future scope

The practical execution of the above-designed model is theoretically proven to be more efficient and most appropriate in the Emergency Department of any hospital. The results clearly show that digital twins have a fast start collaborating with different hospital stakeholders and patients. The future scope of this proposal to design the expert system to decide for generalized emergency examinations required Data Analysis of existing data and, finally, prediction of the diagnosis needed. The recommendation of the treatment plan for specific diagnosis.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] S. Zeb, A. Mahmood, S.A. Hassan, M.D. Jalil Piran, M. Gidlund, M. Guizani, Industrial digital twins at the nexus of nextG wireless networks and computational intelligence: a survey, J. Netw. Comput. Appl. (2022), 103309.
[2] R. Sahal, S.H. Alsamhi, K.N. Brown, D. O'Shea, B. Alouffi, Blockchain-based digital twins collaboration for smart pandemic alerting: decentralized COVID-19 pandemic alerting use case, Computat. Intell. Neurosci. (2022). *2022*.

[3] E. Henrichs, T. Noack, A.M. Pinzon Piedrahita, M.A. Salem, J. Stolz, C. Krupitzer, Can a byte improve our bite? An analysis of digital twins in the food industry, Sensors 22 (1) (2022) 115.
[4] J. Zhang, J.C. Cheng, W. Chen, K. &Chen, Digital twins for construction sites: concepts, LoD definition, and applications, J. Manage. Eng. 38 (2) (2022), 04021094.
[5] Z. Lv, J. Guo, H. Lv, Safety poka-yoke in zero-defect manufacturing based on digital twins, IEEE Trans. Ind. Inf. (2022).
[6] X. Liu, J. Jiang, Digital Twins by Actual Education Teaching Practice in Visual Sensing Training System, Advances in Civil Engineering, 2022, p. 2022.
[7] P. Angin, M.H. Anisi, F. Göksel, C. Gürsoy, A. Büyükgülcü, Aguilera: a digital twin framework for smart agriculture, J. Wirel. Mob. Networks Ubiquitous Comput. Depend. Appl. 11 (4) (2020) 77–96.
[8] S. Neethirajan, B. Kemp, Digital twins in livestock farming, Animals 11 (4) (2021) 1008.
[9] F. Psarommatis, G. May, A literature review and design methodology for digital twins in the era of zero defect manufacturing, Int. J. Prod. Res. (2022) 1–21.
[10] O. Veledar, V. Damjanovic-Behrendt, G. Macher, Digital twins for dependability improvement of autonomous driving, in: Proceedings of the European Conference on Software Process Improvement, Cham, Springer, 2019, pp. 415–426.
[11] Y. Liu, L. Zhang, Y. Yang, L. Zhou, L. Ren, F. Wang, M.J. Deen, A novel cloud-based framework for the elderly healthcare services using digital Twin, IEEE Access 7 (2019) 49088–49101.
[12] M. Braun, Represent me: please! towards ethics of digital twins in medicine, J. Med. Ethics 47 (6) (2021) 394–400.
[13] K. Bruynseels, F. Santoni de Sio, J. van den Hoven, Digital twins in health care: ethical implications of an emerging engineering paradigm, Front. Genet. 9 (2018) 31.
[14] M.N. Kamel Boulos, P. Zhang, Digital twins: from personalized medicine to precision public health, J. Pers. Med. 11 (8) (2021) 745.
[15] S. Elkefi, O. Asan, Digital Twins for Managing Health Care Systems: rapid literature review, J. Med. Internet Res. 24 (8) (2022), e37641.
[16] D.V. Dimitrov, Medical internet of things and big data in healthcare, Healthc. Inform. Res. 22 (3) (2016) 156–163.
[17] U. Sivarajah, M.M. Kamal, Z. Irani, V. Weerakkody, Critical analysis of big data challenges and analytical methods, J. Bus. Res. 70 (2017) 263–286.
[18] B. Kajwang, Implications for big data analytics on claims fraud management in insurance sector, Int. J. Technol. Syst. 7 (1) (2022) 60–71.
[19] J. Corral-Acero, F. Margara, M. Marciniak, C. Rodero, F. Loncaric, Y. Feng, P. Lamata, The 'digital Twin'to enable the vision of precision cardiology, Eur. Heart J. 41 (48) (2020) 4556–4564.
[20] Scoles, S. (2016). A digital twin of your body could become a critical part of your health care. *Available online*: https://slate.com/technology/2016/02/assaults-living-heart-project-and-the-future-of-digital-twins-in health-care. *HTML (accessed on 20 June)*.
[21] K. Gillette, M.A. Gsell, A.J. Prassl, E. Karabelas, U. Reiter, G. Reiter, G. Plank, A Framework for the generation of digital twins of cardiac electrophysiology from clinical 12-leads ECGs, Med. Image Anal. 71 (2021), 102080.
[22] R.M. Portela, C. Varsakelis, A. Richelle, N. Giannelos, J. Pence, S. Dessoy, M. von Stosch, When Is an In Silico Representation a Digital Twin? A Biopharmaceutical Industry Approach to the Digital Twin Concept, 2020.
[23] M. Canzoneri, A. De Luca, J. Harttung, Digital Twins: A General Overview of the Biopharma Industry, 177, Digital Twins: Applications to the Design and Optimization of Bioprocesses, 2021, pp. 167–184.
[24] K. Subramanian, Digital twin for drug discovery and development – the virtual liver, J. Indian Inst. Sci. (2020) 1–10.
[25] F. Tao, Q. Qi, L. Wang, A.Y.C. Nee, Digital twins and cyber–actual systems toward smart manufacturing and industry 4.0: correlation and comparison, Engineering 5 (4) (2019) 653–661.
[26] O. Mazumder, D. Roy, S. Bhattacharya, A. Sinha, A. Pal, Synthetic PPG generation from the haemodynamic model with baroreflex autoregulation: a digital twin of the cardiovascular system, in: Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2019, pp. 5024–5029.
[27] A. Croatti, M. Gabellini, S. Montagna, A. Ricci, On the integration of agents and digital twins in healthcare, J. Med. Syst. 44 (9) (2020) 1–8.
[28] S. Montagna, A. Croatti, A. Ricci, V. Agnoletti, V. Albarello, E. Gamberini, Real-time tracking and documentation in trauma management, Health Inf. J. 26 (1) (2020) 328–341.
[29] S. Montagna, A. Omicini, Agent-based modelling for the self-management of chronic diseases: an exploratory study, Simulation 93 (9) (2017) 781–793.
[30] S. Vairavasundaram, V. Varadarajan, D. Srinivasan, V. Balaganesh, S.B. Damerla, B. Swaminathan, L. Ravi, Dynamic physical activity recommendation delivered through a mobile fitness app: a deep learning approach, Axioms 11 (7) (2022) 346.
[31] N.H. Strickland, PACS (picture archiving and communication systems): filmless radiology, Arch. Dis. Child. 83 (1) (2000) 82–86.
[32] J. Sligo, R. Gauld, V. Roberts, L. Villa, A literature review for large-scale health information system project planning, implementation and evaluation, Int J Med Inform 97 (2017) 86–97.
[33] J. Wang, Patient flow modeling and optimal staffing for emergency departments: a petri net approach, IEEE Trans. Comput. Soc. Syst. (2022).
[34] D. AOCNP, The evolution of the electronic health record, Clin. J. Oncol. Nurs. 19 (2) (2015) 153.
[35] Z. Wang, Z. Yang, T. Dong, A review of wearable technologies for elderly care that can accurately track indoor position, recognize actual activities and monitor vital signs in real-time, Sensors 17 (2) (2017) 341.
[36] S. Tian, W. Yang, J.M. Le Grange, P. Wang, W. Huang, Z. Ye, Smart healthcare: making medical care more intelligent, Glob. Health J. 3 (3) (2019) 62–65.
[37] C. Herwig, R. Pörtner, J. Möller, eds.. Digital Twins: Tools and Concepts for Smart Biomanufacturing, Springer International Publishing, Cham, 2021.
[38] M. Alazab, LU. Khan, S. Koppu, S.P. Ramu, M. Iyapparaja, P. Boobalan, T. Baker, P. K.R. Maddikunta, T.R. Gadekallu, A. Aljuhani, Digital twins for healthcare 4.0-recent advances, architecture, and open challenges, IEEE Consumer Electron. Mag. (2022).
[39] A. Durojaiye, J. Fackler, N. McGeorge, K. Webster, H. Kharrazi, A. Gurses, Examining diurnal differences in multidisciplinary care teams at a pediatric trauma center using electronic health record data: social network analysis, J. Med. Internet Res. 24 (2) (2022) e30351.
[40] S. Nam, D. Kim, W. Jung, Y. Zhu, Understanding the research landscape of deep learning in biomedical science: scientometric analysis, J. Med. Internet Res. 24 (4) (2022) e28114.
[41] A. Esteva, B. Kuprel, R.A. Novoa, J. Ko, S.M. Swetter, H.M. Blau, S. Thrun, Dermatologist-level classification of skin cancer with deep neural networks, Nature 542 (7639) (2017) 115–118.
[42] Lu Wang, H. Wang, Y. Huang, B. Yan, Z. Chang, Z. Liu, M. Zhao, L. Cui, J. Song, F. Li, Trends in the application of deep learning networks in medical image analysis: evolution between 2012 and 2020, Eur. J. Radiol. 146 (2022), 110069.
[43] A. Kumar, P. Tejaswini, O. Nayak, A.D. Kujur, R. Gupta, A. Rajanand, M. Sahu, A survey on IBM watson and its services, in: Proceedings of the Journal of Physics: Conference Series 2273, IOP Publishing, 2022, 012022.
[44] S.F. Merck, Chronic disease and mobile technology: an innovative tool for clinicians, Nurs. Forum 52 (4) (2017) 298–305.
[45] Almulhim, A., Hartley H., Norman P., Caton S., and Goyder E. "Behavioural change techniques in health coaching-based interventions for type 2 diabetes patients: a systematic review." (2022).
[46] S.M. Schwartz, K. Wildenhaus, A. Bucher, B. Byrd, Digital twins and the emerging science of self: implications for digital health experience design and "small" data, Front. Comput. Sci. 2 (2020) 31.
[47] T.M. Morrison, P. Pathmanathan, M. Adwan, E. Margerrison, Advancing regulatory science with computational modeling for medical devices at the FDA's office of science and engineering laboratories, Front. Med. (2018).
[48] F. Pappalardo, G. Russo, F.M. Tshinanu, M. Viceconti, Silico Clinical trials: Concepts and Early Adoptions, Brief Bioinform, 2019.
[49] J. Andreu-Perez, D.R. Leff, H.M. Ip, G.Z Yang, From wearable sensors to smart implants-–toward pervasive and personalized healthcare, IEEE Trans. Biomed. Eng. 62 (12) (2015) 2750–2762.
[50] N. Reddy, N. Verma, K. Dungan, Monitoring Technologies-Continuous Glucose Monitoring, Mobile Technology, Biomarkers of Glycemic Control, Endotext [Internet], 2020.
[51] D. Zhang, Q. Liu, Biosensors and bioelectronics on smartphone for portable biochemical detection, Biosens. Bioelectron. 75 (2016) 273–284.
[52] A.O. Akmandor, N.K. Jha, Keep the stress away with SoDA: stress detection and alleviation system, IEEE Trans. Multi-Scale Comput. Syst. 3 (4) (2017) 269–282.
[53] M. Chan, E. Campo, D. Estève, J.Y Fourniols, Smart homes—current features and future perspectives, Maturitas 64 (2) (2009) 90–97.
[54] L. Liu, E. Stroulia, I. Nikolaidis, A. Miguel-Cruz, A.R. Rincon, Smart homes and home health monitoring technologies for older adults: a systematic review, Int. J. Med. Inform. 91 (2016) 44–59.
[55] H. Yin, N.K. Jha, A health decision support system for disease diagnosis based on wearable medical sensors and machine learning ensembles, IEEE Trans. Multi-Scale Comput. Syst. 3 (4) (2017) 228–241.
[56] A. Croatti, M. Gabellini, S. Montagna, A. Ricci, On the integration of agents and digital twins in healthcare, J. Med. Syst. 44 (9) (2020) 1–8.
[57] D. Estrin, I. Sim, Open mHealth architecture: an engine for health care innovation, Science 330 (6005) (2010) 759–760.
[58] M.P. Gagnon, P. Ngangue, J. Payne-Gagnon, M. Desmartis, m-Health adoption by healthcare professionals: a systematic review, J. Am. Med. Inform. Assoc. 23 (1) (2016) 212–220.
[59] N. Bakkar, T. Kovalik, I. Lorenzini, S. Spangler, A. Lacoste, K. Sponaugle, R. Bowser, Artificial intelligence in neurodegenerative disease research: use of IBM Watson to identify additional RNA-binding proteins altered in amyotrophic lateral sclerosis, Acta Neuropathol. 135 (2) (2018) 227–247.
[60] N. Renugadevi, S. Saravanan, C.N. Sudha, Revolution of smart healthcare materials in big data analytics, Mater. Today Proc. (2021).
[61] J. Kanevsky, J. Corban, R. Gaster, A. Kanevsky, S. Lin, M. Gilardino, Big data and machine learning in plastic Surgery: a new frontier in surgical innovation, Plast. Reconstr. Surg. 137 (5) (2016) 890e–897e.
[62] J.R. Díaz-Palacios, V.J. Romo-Aledo, A.H. Chinaei, Biometric access control for e-health records in pre-hospital care, in: Proceedings of the Joint EDBT/ICDT 2013 workshops, 2013, pp. 169–173.
[63] D. Bhattacharyya, R. Ranjan, F. Alisherov, M. Choi, Biometric authentication: a review, Int. J. u-e-Serv. Sci. Technol. 2 (3) (2009) 13–28.
[64] H. Al-Assam, W. Hassan, S. Zeadally, Automated biometric authentication with cloud computing. Biometric-Based Actual and Cybersecurity Systems, Springer, Cham, 2019, pp. 455–475.
[65] T.M. Morrison, P. Pathmanathan, M. Adwan, E. Margerrison, Advancing regulatory science with computational modeling for medical devices at the FDA's Office of Science and Engineering Laboratories, Front. Med. (Lausanne) 5 (2018) 241.
[66] F. Pappalardo, G. Russo, F.M. Tshinanu, M. Viceconti, In silico clinical trials: concepts and early adoptions, Brief. Bioinform. 20 (5) (2019) 1699–1708.

[67] A.E. Onile, R. Machlev, E. Petlenkov, Y. Levron, J. Belikov, Uses of the digital twins concept for energy services, intelligent recommendation systems, and demand-side management: a review, Energy Rep. 7 (2021) 997–1015.
[68] C. Rajeswari, B. Sathiyabhama, S. Devendiran, K. Manivannan, Bearing fault diagnosis using wavelet packet transform, hybrid PSO and support vector machine, Proced. Eng. 97 (2014) 1772–1783.
[69] C. Rajeswari, B. Sathiyabhama, S. Devendiran, K. Manivannan, A gear fault identification using wavelet transform, rough set based GA, ANN and C4. 5 algorithm, Proced. Eng. 97 (2014) 1831–1841.
[70] N.Y. Isabelle, mMES: a mobile medical expert system for health institutions in Ghana, Int. J. Sci. Technol. 2 (6) (2012) 333–344.
[71] M. Karabatak, M.C. Ince, An expert system for detection of breast cancer based on association rules and neural network, Expert Syst. Appl. 36 (2) (2009) 3465–3469.
[72] H.R. Hasan, K. Salah, R. Jayaraman, M. Omar, I. Yaqoob, S. Pesic, D. Boscovic, A blockchain-based approach for the creation of digital twins, IEEE Access 8 (2020) 34113–34126.
[73] R. Jaichandran, Biometric based user authentication and privacy-preserving in cloud environment, Turk. J. Comput. Math. Educ. (TURCOMAT) 12 (2) (2021) 347–350.
[74] A. Holzinger, A. Saranti, C. Molnar, P. Biecek, W. Samek, Explainable AI methods-a brief overview. International Workshop on Extending Explainable AI Beyond Deep Models and Classifiers, Springer, Cham, 2022, pp. 13–38.
[75] S. Whitelaw, D.M. Pellegrini, M.A. Mamas, M. Cowie, H.G.C.V. Spall, Barriers and facilitators of the uptake of digital health technology in cardiovascular care: a systematic scoping review, Eur. Heart J. Digit. Health 2 (1) (2021) 62–74.
[76] M.E.M. Manurung, J. Sitorus, A. Hutahaean, The relationship between the level of knowledge of nurses and response time in the emergency installation at porsea regional general hospital toba regency, J. Midwifery Nurs. 4 (2) (2022) 55–62.
[77] M.M. Ilias, K. Santosh Kumar, K. Nandini, Md Arief, M. Rishikesh, K. Kavya, Biometric authentication for cloud services, J Algebr Stat 13 (3) (2022) 2132–2142.

RAJANIKANTH ALUVALU (Senior Member, IEEE) received the Ph.D. degree in cloud computing as specialization. He is currently working as Professor with the Department of IT, Chaitanya Bharathi Institute of Technology, Hyderabad, India. Formerly, he had held positions, including a Professor and the Head of the Department of CSE, Vardhaman College of Engineering, Hyderabad, the Vice-Chair of the Entrepreneurship and Startup Committee, and a Treasurer and the Secretary of the IEEE Computer Society, Hyderabad Section. He is having more than 17 years of teaching experience. He has organized various international conferences and delivered keynote addresses. He has published more than 60 research papers in various peer-reviewed journals and conferences. He is a Life Member of ISTE and a member of ACM and MIR Labs.

He was a recipient of the Best Advisor Award from the IEEE Hyderabad Section as well as the IUCEE Faculty Fellow Award (2018). He has also been the Co-coordinator of the AICTE 'Margadarshan' Scheme, a Google Cloud Facilitator, an Editor of IJDMMM journal published by Inderscience, and the Academic Editor for Computer Science journal (Peerj), in addition to being the reviewer for several Scopus indexed and SCI indexed journals.

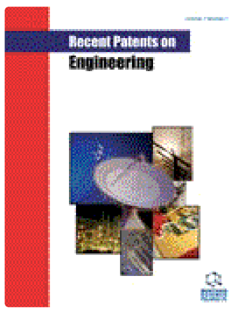
V. UMA MAHESWARI (Member, IEEE) received the Ph.D. degree in image analytics and data science from Visvesvaraya Technological University, Belagavi. She is currently working as an Associate Professor with the Department of CSE, Chaitanya Bharathi Institute of Technology, Hyderabad, India. She has published more than 20 research papers in SCI, ESCI, WoS, DBLP, and SCOPUS indexed journals and conferences. She has also published four Indian patents on facial expression analysis in the fields of medical, e-commerce, education, and security. She has done an enormous study and given contributions in facial expression analysis and applications. She constructed feature vector for a given image based on the directions and introduced dynamic threshold values while comparing the images, which helps to analyze any image. She has researched the similarity of images in a given database to retrieve the relevant images. She also worked with convolutional neural networks by giving the pre-processed input image to improve the accuracy. It has been proved that the maximum edge intensity values are enough to retrieve the required feature from the image instead of working on total image data. She is the Co-ordinator for TEDxVCE. She has organized various technical programs and served as a technical committee member and a reviewer for various conferences. She has delivered sessions in various capacities. She received the Best Faculty Award under the innovation category from the CSI Mumbai Chapter for the year 2019.

Mudrakola Swapna Working as Asst professor at matrusri engineering College of Engineering, Shamshabad, Hyd, T.S. Education Qualification pursuing PhD from Osmania University, research area Artificial Intelligence in the medical application. Completed M.Tech in Information Technology and B.Tech in CSE. Total Teaching Experience is 14 years, and my Industry experience is two years as a software engineer. Moto of my life to empowering myself with my sincere hard work.

Dr. MVS. Sandhya completed her B.Sc in Microbiology from Osmania University in 2009. She received her M.*Sc* in Microbiology from Osmania University in 2011. Holds doctoral degree in Microbiology from Osmania university. Her research work is based on the "Green synthesis of nanoparticles for development of Nanoprobes Assay for Early Diagnosis of Sepsis". Currently working in private sector in *in vitro* assay development department in pre clinical research.

Dr.A.C.Kaladevi working as Professor in the Department of Computer Science and Engineering at Sona College of Technology, Salem, India has more than 25 years of teaching experience. She obtained her B.*Sc* degree in Computer Science from Cauvery College for Women, Tiruchirapalli followed by MCA at PSG College of Technology, Coimbatore. She completed M.Phil Computer Science from Manonmaniam Sundaranar University, Tirunelveli and M.E Computer Science and Engineering from V.M. K.V Engineering College, Salem which was then affiliated to Anna University, Chennai. She was awarded Ph.D degree in Information and Communication Engineering during 2014 by Anna University, Chennai. Her research interest includes Data Analytics, Cloud Computing, and Image Processing. She has published 21 papers in various International Journals and presented 32 papers in both national and international conferences. She has co-authored 3 books in computer science discipline. She has conducted 2 national workshops one on "Big data and Cloud for bigger transformations" funded by Department of Science and Technology (DST), New Delhi under BDI Scheme and the other one on "Empowering the Tribal Women in and around Yercaud Hills, Salem by inculcating self -employment opportunities using innovative ICT based skill development techniques" funded by Tamil Nadu State Council for Science and Technology (TNSCST), Chennai under Dissemination of Innovative Technology Scheme. She has guided more than 25 PG and 40 UG projects out of which 3 UG projects were funded by TNSCST under Students Project Scheme. As an enthusiastic student counselor she has given a great moral support to students who are now placed in a much renowned positions in their career.

Dr. Rohith has completed his Engineering in CSE from Dr. MGR Engineering College, Chennai and completed his Masters and Ph.D from Dr. MGR Educational and Research Institute. His-Passion for Academics and Research has made him the right person to be in Academic Sector. His-areas of Interest include Process Mining, IOT, Machine Learning, Deep Learning and to name a few. His-academic Interest has made him to pursue his research work on Process mining and classification in general. He has got Patents to his credit and had Published Journals in Journals of Repute. Dr. Rohith has got immense interest in Training the students for their Placements and he is currently serving as Professor in Saveetha School of Engineering (SIMATS), Chennai.



Recent Patents on Engineering

Editor-in-Chief

ISSN (Print): 1872-2121
ISSN (Online): 2212-4047

Research Article

A Secure Network with Minimization of Energy for E-healthcare Application in IoMT

In Press, (this is not the final "Version of Record"). Available online 12 June, 2023

Author(s): Rajanikanth Aluvalu, Uma Maheswari V.*, Mohan A. and Yadaiah Balagoni

Published on: 12 June, 2023

Article ID: e280423216332

DOI: 10.2174/1872212118666230428105619

Note! Please note that this article is currently in the "Article in Press" stage and is not the final "Version of record". While it has been accepted, copy-edited, and formatted, however, it is still undergoing proofreading and corrections by the authors. Therefore, the text may still change before the final publication. Although "Articles in Press" may not have all bibliographic details available, the DOI and the year of online publication can still be used to cite them. The article title, DOI, publication year, and author(s) should all be included in the citation format. Once the final "Version of record" becomes available the "Article in Press" will be replaced by that.

Abstract

Aims: Protect patient healthcare records.

Background: The adaptability of the digital healthcare system is a major factor in its recent rise in popularity. Utilizing the digital healthcare system has resulted in an ever-increasing number of healthcare apps. The Internet of Medical Things-(IoMT) is a newly emerging digital healthcare system using various biomedical sensors and the cutting-edge capabilities of wireless systems and cloud computing. Since IoMT can exchange data between various connecting nodes thanks to the combination of other technologies, security and energy consumption provide the greatest challenge to the IoMT infrastructure.

Objective: Reduce the cost of communication in order to strengthen defenses against unauthorized access and increase energy efficiency.

Method: This study provides a protocol for protecting patients; medical records called the requesttype- based energy-aware framework (Re-EAF) based on patent. The primary goal is to reduce the cost of communication in order to strengthen defences against unauthorized access and increase energy efficiency. An identifying unit called a request-type energy aware framework has been proposed. The proposed method avoids treating all requests the same by instead characterizing them based on the identified criteria and characteristics. Using Constrained Application Protocol (CoAP), remote patient monitoring can increase the safety of gathered data.

Results: Using Constrained Application Protocol (CoAP), remote patient monitoring can increase the safety of gathered data. Using a software-defined networking (SDN) framework, our research ensures that data and requests are sent and received as effectively and efficiently as possible while conserving energy.

Conclusion: In this research, the transmitted healthcare data is encrypted via cipher Block-chaining. The experimental study demonstrates that the suggested Re-EAF consumes less energy while producing a higher throughput than conventional methods.

Keywords: Constrained application protocol, block-chain, internet of medical things, request-type based energy aware framework, energy consumption, security, constrained application protocol.

Multimedia Tools and Applications (2023) 82:20989–21004
https://doi.org/10.1007/s11042-023-14460-8

A proficient video recommendation framework using hybrid fuzzy C means clustering and Kullback-Leibler divergence algorithms

H. Anwar Basha[1] · S. K. B Sangeetha[2] · S. Sasikumar[3] · J. Arunnehru[2] · M. Subramaniam[4]

Received: 14 April 2021 / Revised: 9 February 2022 / Accepted: 31 January 2023 /
Published online: 7 February 2023

Abstract

A video recommendation framework for e-commerce clients is proposed using the collaborative filtering (CF) process. One of the most important features of the CF algorithm is its scalability. To avoid the issue, a hybrid model-based collaborative filtering approach is proposed. KL Divergence was developed to address the CF technique's scalability problem. The clustering with enhanced sqrt-cosine similarity Recommender scheme is proposed. For successful clustering, Kullback–Leibler Divergence-based Fuzzy C-Means clustering is suggested, with the aim of focusing on greater accuracy during movie recommendation.The proposed scheme is viewed as a trustworthy contribution that significantly improves the ability of movie recommendation by virtue of the KL divergence-based Fuzzy C-Means clustering mechanism and enhanced sqrt-

✉ H. Anwar Basha
anwar.mtech@gmail.com

S. K. B Sangeetha
sangeets8@srmist.edu.in

S. Sasikumar
sasikumar@saveetha.ac.in

J. Arunnehru
arunnehru.aucse@gmail.com

M. Subramaniam
subbu.21074@gmail.com

[1] School of Computer Science and Engineering, REVA University, Bengaluru, India

[2] Department of Computer Science and Engineering, College of Engineering and Technology, Faculty of Engineering and Technology, SRM Institute of Science and Technology, Vadapalani Campus, Chennai, India

[3] Department of Computer Science and Engineering, Saveetha Engineering College, Chennai, India

[4] Department of CSE , Chaitanya Bharathi Institute of Technology, Hyderabad, India

cosine similarity. The proposed scheme highlighted and addressed the critical role of the KL divergence-based cluster ensemble factor in improving clustering stability and robustness. For prediction, the enhanced sqrt-cosine similarity was used to calculate successful related neighbor users. The performance of Recommendation is improved when KLD-FCM is combined with improved sqrt-cosine similarity.The proposed scheme's empirical work on the Movielens dataset in terms of MAE, RMSE, SD, and Recall were found to be superior in recommendation accuracy compared to traditional approaches and some non-clustering based methods recommended for study. With the specified number of clusters, it is capable of providing accurate and customized movie recommendation systems.

Keywords Recommendation systems · Content-based filtering · Knowledge base · Fuzzy C-means algorithm

1 Introduction

A Recommendation System (RS) is a filtering program that enables consumers to evaluate product recommendations for online purchases and provide information into products that they are interested in. In recent years, the extensive advancement of science and technology has resulted in vast amounts of digital knowledge being accessible via the internet. As a result of the overabundance of data, users are unable to obtain accurate taste information. This problem of information overload can be solved with RS, which filters out irrelevant information and only recommends related things to users.RS's knowledge filtering system assists users in making decisions in difficult situations by scanning vast data for items of interest. It also offers a personalized proposal based on the user's desires and preferences. RS is widely used in e-learning, e-shopping, e-tourism, e-business, e-government, and social networking sites like Facebook. Articles from research, books, music, news, films, DVDs/CDs, and other e-shopping products recommended by RS for their clients [1, 2].

To create a working recommendation system, a large amount of data must be collected. RS accepts a variety of inputs, both explicit and implicit. The explicit ratings of 1 to 5 given by users for their preferences in the items they purchase make up the covert reviews. User actions such as accessing and navigating past websites, click and search logs provide implicit input. Demographic data is another addition to RS. This information index was developed for each client who visits the site. Following this point, the data is filtered to obtain sufficient data for customer/user suggestions.Content-based filtering, collaborative filtering, and hybrid filtering technology are examples of filtering algorithms that would be more suitable for the recommending engine [3, 4]. The collection of data and application of recommendation filtering methods yield a set of recommendations that comply with the procedures to be considered in a recommendation system's calculation. In general, two types of performance are predicted and suggested. The ranking items for which the target consumer has not been rated are forecasted by prediction. On the basis of the forecasted ratings, the recommendation recommends the top-n recommendations to the target consumer, where each item does not include an evaluation value for these top-n recommendations. The consumer should be ecstatic with the performance of the recommender [5].

A framework that offers good and helpful recommendations for its own users requires the use of appropriate and reliable recommendation techniques. The content-based technology employs a domain-based algorithm that focuses on analyzing the characteristics of predictive posts. When documents such as blogs, magazines, and news are recommended, it is the most effective material filtration technique. The user profile recommendation in the Content-Based Filtering (CBF) approach is based on characteristics extracted from the content of items checked by the user.Objects that are mostly associated with positive items are recommended to the customer. CBF employs a variety of models to identify similarities between documents in order to generate useful recommendations. To form the relationship between different documents within the Corpus, it could use a Vector Space Model or a probabilistic model like the Naive Bayes Classification method, Decision Trees, or Neural Networks. It's also possible to use the Vector Classification method. These approaches make recommendations based on the underlying model's statistical analysis or machine learning techniques [6].

Other users' profiles aren't needed for content-based filtering because they don't affect the recommendation. Furthermore, as the user profile changes, CBF's approach adjusts its recommendations in a very short time. The key drawback of this method is that it necessitates thoroughly informing and explaining the characteristics of the objects in the profile. CF is a domain-agnostic content prediction technique that can't be easily or reliably classified by metadata like film or music. Filtering by working together creates a database of user preferences (user-item matrix). It then matches individuals with relevant interests and preferences to make recommendations.This group of people is creating a social network. A consumer receives recommendations for items that he hasn't yet rated but that other users in his field have given high marks to. CF may make recommendations in the form of forecasts or recommendations. Collaborative filtering has a range of benefits over CBF, including the ability to be used in fields where object content is scarce and computer system content is difficult to assess (such as opinions and ideals). CF technology should provide persuasion recommendations so that things that are useful to the user can be recommended, even though the user profile does not include the content [9].

Hybrid filtration strategies combine various recommendation approaches to boost device optimization and prevent some of the drawbacks and issues that come with pure recommending systems. Since one algorithm can overcome its drawbacks with another algorithm, a combination of algorithms is built to make recommendations more accurately and efficiently than a single algorithm. Using various recommendation models, a combined model will eradicate the flaws of a single process.Separate algorithms can be applied and the results merged, content-based co-filters can be used, content-based collaborative filters can be used, and a single recommendation system can be developed to bring all methods together [10].

The most promising products from which consumers can choose are included in the recommendation question. Some well-known approaches for solving the problem of scalability under model-based collaboration filters are clustering-based approaches. Predominantly, the majority of CF-based clustering strategies have relied on K-means and Fuzzy C-means clustering, which lack the ability to pick a relevant clustering core, lowering the predictive efficiency. As a result, there are trade-offs between scalability and predictive efficiency.Improved clustering techniques were suggested in the study to recognize agreed problems due to scalability. Many previous studies have shown that clustering-based CF systems (CF combined with clustering algorithms) are a promising schema for providing accurate personal recommendations and solving large-scale problems [8]. Fuzzy C-Means is a soft clustering approach that allows each individual data to be allocated to multiple clusters based on different membership degrees. They also concluded that good clustering-based CF performance is dependent on appropriate clustering techniques as well as the dataset's design. The analysis shows that the Clustering algorithm's stability and robustness need to be improved in order to achieve critical accuracy in the process of movie recommendation to target users. The following are the limitations of the current method:

- Fuzzy-C refers to a lack of ability to choose the initial cluster Center point, which can lead to a local optimum solution and affect clustering accuracy. In certain cases, the obtained clusters can be impractical, affecting the Recommendation outcomes.
- To get a good grouping of data, most current clustering algorithms require the configuration of some parameters.
- The disadvantage of FCM is that it has a higher error rate and needs further iterations to obtain well-framed clusters.
- Because of the prejudices and assumptions that each clustering algorithm contains, applying a single clustering method generally results in inconsistent results.

The Fuzzy C Means clustering with KL divergence is suggested to solve the aforementioned limitations. The research work's contributions have the following features.

- To prevent the drawbacks of a bad initialization, Ensemble FCM clustering is used to divide users into separate groups.
- The Ensemble Fuzzy C-Means clustering methods use a KL divergence-based cluster ensemble factor to improve the stability and accuracy of the clustering process, resulting in successful clustering with the goal of focusing on better performance results during movie recommendation.
- A better approach is to treat the membership vector as a discrete probability function, with the statistical distance, such as KL divergence, serving as the similarity metric.
- For active users, the enhanced sqrt-cosine similarity is often used to find the most powerful nearest neighbors.
- Reduce the problem of scalability.

The latest analysis of RS methods is summarized in Section 2. In addition, work on various RS is discussed in this chapter. The proposed framework model for the hybrid video recommender is defined in Section 3. The quality of predictions as measured by the assessment metrics is also stated. Section 4 brings the analysis to a close by highlighting the algorithms that aided in the achievement of the objectives and promoted the desired outcome. The study's limitations have been established, and potential research directions have been summarized.

2 Related study

Recommend systems use data mining and predictive algorithms to predict user preferences among the vast array of images, goods, and services available. The rapid growth of knowledge on the Internet, as well as the number of visits to websites, are posing significant challenges to system recommendations. The development of precise recommendations, the successful management of a large number of recommendations, and the large number of system members are all examples of these challenges.As a result, new system recommendation technologies are needed that can produce high-quality recommendations quickly, even for large data sets. There are numerous methods and algorithms for data filtering and recommendation. This section provides a brief overview of recent system-related studies in the literature.

Collaborative filtering is a widely used and relevant technology that makes predictions and suggestions based on the ratings and actions of other system users. The key premise of this strategy is that the user can pick and consolidate views from other users in order to understand his choice for the active user. Memory-based CF algorithms generate a forecast for the entire database or a subset of the user-item database. Each consumer is a member of a group of people who share similar interests. A prediction of a new user's tastes for new objects can be rendered by identifying related neighbors (or active users).Memory-based collective filtering has a number of drawbacks, the most significant of which is that it must use the entire database every time it predicts something, making it extremely sluggish in memory. If the rating matrix is so broad that many people use it, the issue becomes serious. Computing resources are depleted, and device performance suffers as a result, making it impossible for the system to respond quickly to user requests [11].

The model-based approach learns a model to boost collaborative filtering technology efficiency using prior scores. The model could be built using machine learning or data mining techniques. Since these approaches rely on pre-computed models, they can quickly suggest a large number of items and have been shown to yield results that are comparable to neighborhood-based recommendation techniques. Dimension reduction, for example, includes techniques including singular decomposition, matrix completion, latent semantic approaches, regression, and clustering.Content-based filters recommend items on a user's item profile and user profile. When an account is created and the framework is first used, these types of profiles are created. As a result of the user's interaction with the system, a better user profile is developed. If a user likes an object in the past, the CBF scheme assumes that the user would like similar things in the future [7]. The most powerful filtering technique is used in information documents such as web sites, journals, and news. To produce meaningful recommendations, CBF employs a variety of models to detect correlations between documents. Model-based vector space models can be used to represent the relationships between various documents within a corpus, such as reverse frequency or probabilistic models like the Naive Bayes Classification, Decision Trees, or Networks. Object metadata is used in the filtration mechanisms. Before users can get a recommendation, they'll need a large collection of items and a well-organized user profile. As a result, the effectiveness of CBF is dependent on the availability of descriptive data. Over-specialization is another major problem with the CBF methodology. Users can only get suggestions that are close to their own [12].

In certain applications, hybrids of various types outperformed individual algorithms. When the algorithms in question cover a wide range of use cases or aspects of the data set, hybrids can be particularly useful. The suggestion has been suggested to be implemented using a range of approaches, including material-, collaborative-, knowledge-based, and other techniques.Every form of recommendation has its own set of strengths and weaknesses. In order to improve efficiency, these strategies were often combined into hybrid recommenders. The hybrid recommender method has a higher level of complexity and implementation costs [14].

The Knowledge Base (KB) suggestion suggests things based on user experience, artifacts, and/or user relationships. In most cases, KB recommendations maintain a knowledge base describing how a particular item serves the needs of a specific person, which can be carried out based on inferences about the relationship between a user's need and a potential recommendation. The semantic similarity between objects can be calculated using the domain ontology. Social recommendation services are an integral part of everyday life on social media.Every minute, users on social media exchange details. Social advisor programs are designed to help people understand what they really want by reducing the amount of information available on social media. They want to help people on social networking sites like Facebook, Twitter, YouTube, Flickr, and Weibo by providing them with tweets and profiles that meet their needs [13].

Systems that are recommended are those that can analyze past user habits and make suggestions for current issues. Simply put, information on similar behavior, remarks, and users will be used in the RSs to try to define the user's thought style in order to assess and suggest the user's taste as the most appropriate and near object. Many of the methods in the RSs are used to make recommendations that are as accurate as the users need.As a result, several models for RSs clustering existing data have been presented in order to efficiently process data with large volumes. Given the goals, special characteristics, and relationships between data in the RSs, using an effective data cluster approach to efficiently process data in subsequent steps and produce more reliable suggestions has always been seen as an important research area for developing these systems [15].

Collaborative filtering recommendation systems are extremely useful for a wide range of online activities, including e-commerce. However, there are significant issues, especially in scalable and dynamic scenario implementations where new users, objects, and ratings are added frequently. Scalability refers to a system's, network's, or process's ability to control or expand its capacity to accommodate development. For example, a device is called escalable if it can increase its total power under increased load as resources (typically hardware) are added.Dimension reduction-based approaches address scalability problems such as SVD, MF, clustering, and so on. In short, cluster-based approaches retain the advantages of low computation cost (for searching candidates) over memory-based approaches as models for dimension reduction [16, 17].

To improve the performance of the recommendation, related users are grouped together based on their interests. Clustering was used to quickly locate a user's neighbors. By reducing the size of the original data to more manageable partitions, clustering systems can react quickly. Clustering, in particular, boosts the scalability and accuracy of recommendation systems. Despite the advantages of low machine costs (for searching candidates) over memory-based and SVD methods, MF methods remain cluster-based methods.One of the most serious problems with a recommending method is scarring, and data sparing has an effect on the accuracy of the recommendation. In general, machine data like Movie Lentils is interpreted as a user-item matrix made up of films, which increases matrix dimensions and sparsity since the user and items are no longer used. Most users do not rate most items, and there are few available ratings. The key explanation for this is a lack of knowledge. In order to condense the user's object matrix, the reduction of dimension addresses the issue of scarcity by excluding non-representative or insignificant users or objects. However, during the reduction process, potentially valuable information is lost.However, some potentially valuable information can be lost during the reduction process [18, 19].

Collaborative filters create this problem because they depend on the rating matrix in most cases. Many researchers have attempted to address this problem, but more research is still required in this area. Most recommendation systems on the major electronic commerce platforms have been influenced by the long tail effect in some way. Since accuracy-focused recommender systems tend to recommend common goods, recommending items with few ratings is critical (long tail items). Popular products that are likely to be less helpful to users can be easily recommended using detailed recommendation algorithms.To assess the ability of systems to recommend unpopular goods, the assessment metrics diversity and innovation have been added. Recommending long tail artifacts can result in the recommendation's precise results being lost. As a result, a recommendation process must be developed that recommends controversial products while minimizing accuracy loss. Several guidelines have recently been proposed to strike a balance between precision, diversity, and novelty [20–22].

Existing CF clustering method algorithms are ineffective at improving RS efficiency and addressing scalability issues. The recommended performance has an effect on the efficiency of the clustering procedures [23, 24]. As a result, there is a lack of precision and coverage, which makes clustering-based approaches in recommender systems difficult to use in practice. To improve the recommendation's performance, better methods for optimizing the above problem are needed.

3 Proposed system model

3.1 KL divergence based ensemble fuzzy C means clustering

Cluster-based CF has been shown to address scalability issues while also improving the consistency of recommendation outcomes in recent years. The aim of clustering algorithms is to group objects into clusters with the shortest distance between them in order to find objects that are identical. Clustering strategies will typically group a large number of users into various clusters based on their rating similarity in order to locate "like-minded" neighbors.One of the most widely used clustering methods is fuzzy clustering. To get a decent grouping of data, most current FCM algorithms require the specification of some parameters. As a result, the Fuzzy cluster ensemble solution usually prevents the drawbacks of a bad initialization.

3.1.1 Fuzzy C means clustering algorithm

1. Initialize Membership matrix M with random data points.
2. Fuzzy cluster center is calculated C
3. Calculate the objective function F = M*(Xi−Cj)
4. For every iteration fuzzy Membership is updated by using M=Σ CK Where k is the number of clusters
5. The iteration will stop when (k+1) − (k) < termination criterion

Ensemble clustering blends a dataset's various simple partitions into a more stable and robust one. The basic concept behind the Ensemble Fuzzy C Means cluster method is to apply the clustering method to the data several times (rather than only once) and then merge the results into a single partition. Ensemble clustering takes a collection of data partitions as input. The cluster ensembles are divided into two parts. The ensemble clustering generator is one, and the consensus function is another.The first section focuses on generating more diverse clustering results, while the second section focuses on seeking a good consensus feature to increase the results' accuracy. Homogeneous ensemble FCM clustering is used in the first component of cluster ensembles. The term "homogeneous ensemble" refers to the use of multiple runs of a single clustering algorithm (fuzzy c-means algorithm) with different initializations and fuzzy parameter values. Several soft partitions of the data are obtained at the end of the first stage of the ensemble method as a result of several runs of the algorithm(s). The aim of this is to improve the accuracy and consistency of fuzzy cluster analysis procedures. Soft ensembles are characterized by the concatenation of membership probability distributions in the second part of cluster ensembles. The obtained partition will be combined in the second stage using a KL divergence-based objective function to produce a single final partition. The Kullback–Leibler (KL) divergence was then used to describe a distance measure between two instances. The similarity of a membership vector to a cluster center is measured by FCM using squared Euclidean distance.This is ineffective in situations where a data's membership in all clusters normally equals one. A better approach is to treat the membership vector as a discrete probability function, with the statistical distance, such as KL divergence, serving as the similarity metric. This algorithm is identical to fuzzy c-means, with the exception that it employs the KL divergence to treat memberships as discrete probabilities.

The data from the User Item is first categorized using homogeneous fuzzy clustering methods. After that, a fuzzy KL divergence-based objective function aggregates the soft clustering effects.

3.2 Improved Sqrt-cosine similarity

Improved Sqrt-Cosine Similarity (ISC) is a modern similarity measure that uses Hellinger distance and is based on sqrt cosine similarity. Hellinger distance (L1 norm) is a much better metric for high-dimensional data mining applications than Euclidean distance (L2 norm). In terms of implementation, the ISC is very similar to cosine similarity, and it outperforms other similarity measures in high-dimensional results.The enhanced sqrt-cosine similarity determines how close two users are. For High Dimensional data, the Hellinger distance-based Similarity is more accurate. The KLD-FCM with enhanced sqrt-cosine similarity outperforms current systems in terms of recommendation performance.

3.3 Proposed KLD-FCM based movie recommendation scheme

For improving movie recommendation methods (KLD FCM-RS), a kullback–leibler divergence-based fuzzy c-means clustering is proposed, and the steps involved in KLD FCM-RS are discussed in detail. The aim of the proposed method is to develop a Collaborative Movie Recommender framework that can solve scalability problems while also improving prediction accuracy in terms of MAE, Precision, Recall, and Speed. The proposed KLD-FCM based Movie Recommendation Scheme is architecturally depicted in Fig. 1 as follows.

It is divided into two phases: offline and online. The User Item Rating Matrix derived from the used Movie Lens Dataset is used as a possible input during the offline process. Then, over the extracted user Item scores, a method of different homogeneous Fuzzy C means clustering is applied to divide the users into different classes. Furthermore, KL Divergence-based cluster ensemble FCM is used to combine the various FCM clustering findings in order to generate efficient single User clusters. The nearest cluster estimation for Active consumer is computed using the Euclidean distance method in the online process.The active user's nearest neighbors in his or her nearest cluster are then found using an enhanced sqrt-cosine similarity tool. Finally, the top list of recommended movie items is calculated in an online mode by determining the movie items that are most frequently recommended by the context's neighborhood users. Leibler–Kullback For the Recommender method, divergence-based Fuzzy C-Means clustering with enhanced sqrt-cosine similarity worked well. Three possible steps are included in this proposed KLD-FCM:

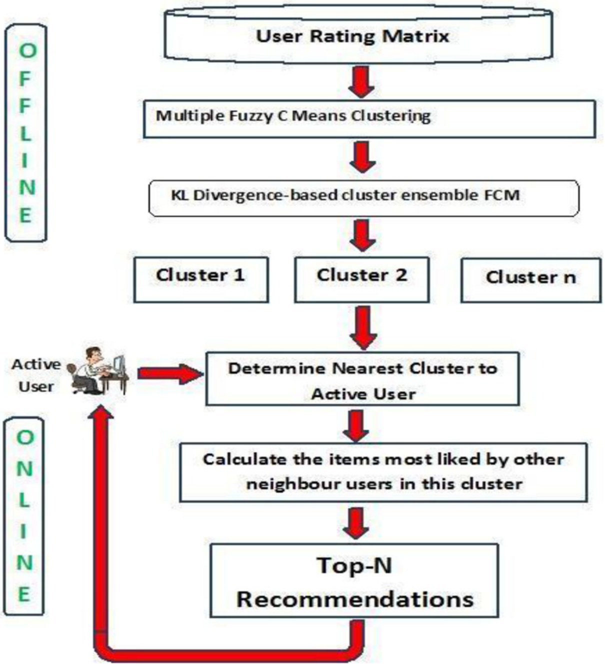


Fig. 1 Architectural view of the proposed KLD-FCM Scheme

3.3.1 Procedure

Step 1 (FCM-KLD) FCM-KLD is divided into two phases. Multiple Fuzzy C Means Clustering is the first step. The user data ratings from the Movie Lens data set are used as input in this phase of the clustering process. Over the User Item scores, apply three different homogeneous Fuzzy C Means clustering methods with different initializations. By executing the FCM several times for each initialization with different fuzzy parameter values, homogeneous FCM clustering is used to create several partitions with different random initializations (here 1.5, 2 and 2.5 are used).**Input:** User Rating Matrix **Output:** Three Clustering results.

Table 1 shows a snapshot of the User object rating matrix from the Movie Lens dataset for 5 users on 5 movies, where U1-U5 are users and M1-M5 are objects (movies). The value 1–5 represents the user's likelihood rating for a specific film. The value '0' denotes that the consumer has not rated (or seen) the film. The recommender framework identifies unrated values and suggests the top N films to the consumer Tables 2, 3, 4, and 5.

Step 2 (determine the nearest cluster to active user) After clustering the users into various clusters, the Euclidean distance approach is used to determine the nearest cluster to Active User.

$$sim_i(Cent_i, U) = \sum_{j=1}^{d} \left(Cent_{i,j} - U_j\right)^2 \tag{1}$$

$Cent_i$ is the centroid of 'i' th cluster, U is the Active User Profile.
d is the dimension of data(Number of Attribute).
$Cent_{i,j}$ is the jth attribute of centroid profile in cluster i.

Table 1 Example of rating matrix from Movie Lens dataset

	M1	M2	M3	M4	……	M_m
U1	5	3	4	3	……	3
U2	4	?	?	?	……	1
U3	?	?	?	?	……	0
U4	5	?	?	2	……	?
….	……	……	……	……	……	……
U_n	4	3	?	?	……	1

Table 2 MAE comparison analysis

Cluster size	5	10	15	20	25
Proposed	1.2	1.2	1.15	1.15	1.14
FCM	1.3	1.29	1.28	1.28	1.27
FCMBAT	1.4	1.39	1.38	1.38	1.37
COA	1.5	1.49	1.48	1.48	1.49

Step 3 (using improved sqrt-cosine similarity, determine the top N recommended movies) Improved sqrt-cosine similarity is used to calculate Active User's nearest neighbors. The following formula is used to determine how close users u1 and u2 are.

$$sim(u1, u2) = \frac{\sum_{i=1}^{m} \sqrt{R_{u1,i} R_{u2,i}}}{\sqrt{\left(\sum_{i=1}^{m} R_{u1,i}\right)} \sqrt{\left(\sum_{i=1}^{m} R_{u2,i}\right)}} \tag{2}$$

m Set of common items rated by user u1 and user u2.
$R_{u1,\,i}$ is the rating given to item 'i' by user u1.
$R_{u2,\,i}$ is the rating given to item 'i' by user u2.

Table 3 RMSE comparison analysis

Cluster Size	5	10	15	20	25
Proposed	0.78	0.77	0.76	0.76	0.75
FCM	0.81	0.8	0.79	0.78	0.77
FCMBAT	0.84	0.84	0.84	0.84	0.81
COA	0.87	0.86	0.86	0.85	0.83

Table 4 Recall analysis

Cluster size	5	10	15	20	25
Proposed	25.28	25.68	27.24	27.8	28.56
FCM	23.12	24.68	25.45	25.12	25.89
FCMBAT	20.12	22.68	23.45	23.12	24.89
COA	18.9	19.89	20.35	20.02	21.67

Table 5 Accuracy analysis

Cluster size	5	10	15	20	25
Proposed	89.3	88.2	88.2	87.1	87
FCM	86.3	86.6	86.3	85.4	85
FCMBAT	84.8	84	83.7	83.5	84.3
COA	82.5	82.4	82.3	82.1	81.67

The Hellinger distance is used to compute the similarity between two vectors in the enhanced sqrt-cosine similarity. This phase is essential for comparing each individual user's rating to the ratings of other users in the clusters. Finally, the top list of recommended movie items that could be suggested to an active user at any time is calculated based on the movie items that are most often recommended by the context's neighborhood users. The movies are recommended to target users who are most likely used by other neighbor users and are not used by him/her during the Recommendation process.The weighted average of the ratings of items in the same cluster neighbor's is used to predict the ranking of unrated items for active users, and then the top-N suggestions list is sent to the active user. The rating of unrated movie (item) 'i' for an active user 'a' is predicted by $P_a(i)$

$$Pa(i) = \underline{R}_a + \frac{\sum_{N\varepsilon Cx} Sim(a,N) \times \left(R_N\ (i) - \underline{R}_N\right.}{\sum_{N\varepsilon Cx}(|Sim\ (a,N)|)} \tag{3}$$

Where,

a	Active User.
$\underline{R}_a$	Average of active user a.
Cx	set of nearest neighbors of active user a belonging to one common cluster.
$\underline{R}_N$	Average rating score given by active user's neighbor N.
$Sim(a,\ N)$	similarity between active user a and Neighbor.

4 Experimental design

Experiments using the publicly accessible Movielens dataset are used to equate the performance of the proposed KLD-FCM with that of baseline recommendation system The Movielens data set used to equate the proposed KLD-FCM scheme to the compared COA, FCM-BAT, FCM, and ICF contains 10,00,000 ratings, with 850 users theoretically rating them. This Movielens data set contains reviews ranging from approximately 1000 to 1513 movies, each scored on a scale of 1 to 5. By partitioning the entire Movielens data set using the k-cross validation process, the performance of the proposed KLD-FCM approach is investigated. The findings were evaluated using 5-fold cross validation. The original dataset is divided into five equal subsets. One is used as a test set (20%), while the other is used as a training set (80%). The procedure is repeated five times, with a different test set selected each time, and the average results recorded.In terms of MAE and RMSE, the proposed method was compared to non-clustering methods such as Basic CF (BCF), User-Based CF (UBCF), SVDM (a variant of Single Value Decomposition (SVD) that uses batch learning with a learning momentum), and RSVD (Regularized SVD model). Various collaborative approaches are selected from the literature to verify the proposed method's role in comparison to other cluster-based techniques.

ICF(integrated Collaborative Framework) The merits of the item k-NN algorithm were used to propose an Integrated Collaborative Framework (ICF). This ICF also provided classification restrictions, ensuring that only potential rules are used during the collaborative filter-based categorization of user ratings.

UPCC (user based CF Pearson correlation coefficient based CF) For the Collaborative filtering recommender scheme, the Pearson Correlation Coefficient test is used to determine how closely two users are related.

FCM (fuzzy C means) The performance of User-Based Collaborative Filtering with Fuzzy C Means is compared to that of other clustering methods such as K-means and Self-Organizing Maps (SOM).

FCMBAT The Fuzzy C-Means and BAT-based Movie Recommendation Scheme (FCM-BAT) is an integrated Fuzzy C-Means and BAT-based Movie Recommendation Scheme for promoting efficient and collaborative recommendation to the target users. This FCM-BAT was proposed to address scalability issues and improve the clustering process, with the aim of improving the consistency of the recommendation process.

COA(cuckoo Optimization Algorithms) Furthermore, a possible movie recommendation system based on k-means and COA is proposed in order to improve the rate of recommendation accuracy when using the Movielens dataset.

4.1 Mean absolute error

The suggested method used to measure MAE can be seen in Fig. 2 for different numbers of neighbors.

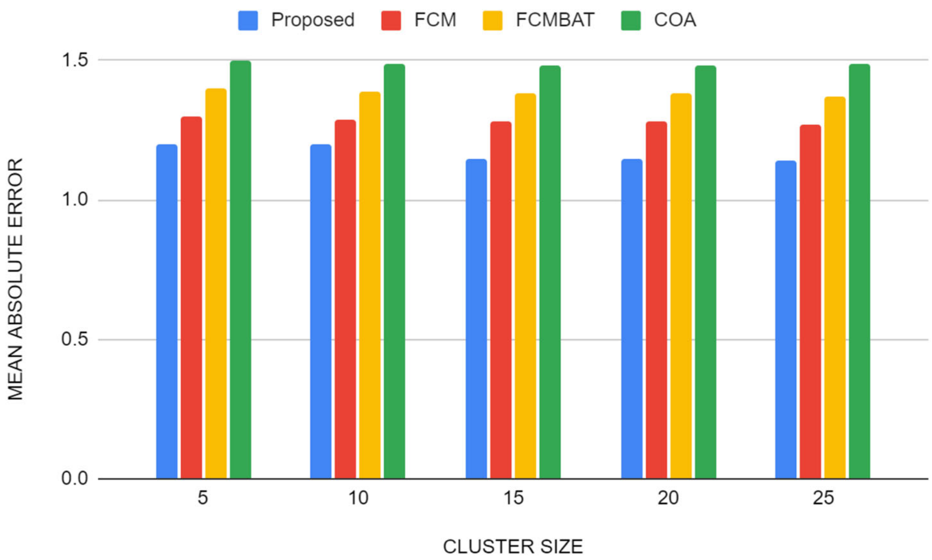


Fig. 2 MAE comparison analysis

Figures 2 and 3 show the contrast of the proposed FCM KLD with current Collaborative Recommender framework methods in terms of MAE and RMSE for different numbers of clusters. The proposed scheme's MAE and RMSE are proving to be significantly lower than those of current schemes. As a result, the proposed scheme's MAE is lower than the COA, FCM-BAT, FCM, and ICF approaches. The proposed scheme's RMSE is also tested to be substantially lower than the baseline schemes under consideration.

Figure 4 graphically depicts the contrast of the proposed KLD-FCM with current Collaborative Recommender framework approaches in terms of recall. It highlights the efficiency of the proposed KLD-FCM scheme as measured by Recall for a variety of cluster sizes. The proposed KLD-FCM scheme's Recall value is determined to be excellent as compared to baseline methods, since the KL divergence factor's guidance in the clustering phase is responsible for the majority of progress.

Figure 5 depicts a graphic comparison of the proposed KLD-FCM with current Collaborative Recommender framework approaches in terms of accuracy. The performance of the proposed method is significantly better than that of current methods. Since it takes advantage of the advantages of KL Divergence Fuzzy C, Clustering with enhanced sqrt-cosine similarity is possible.

5 Conclusion

The KL Divergence Fuzzy C means Clustering with improved sqrt-cosine similarity Recommender framework (KLD-FCM) is proposed to solve the CF technique's scalability problem. For successful clustering, Kullback–Leibler Divergence-based Fuzzy C-Means clustering is suggested, with the aim of focusing on greater accuracy during movie recommendation. The proposed KLD-FCM scheme is described as a trustworthy contribution that significantly improves the ability of movie recommendation by virtue of the KL divergence dependent Fuzzy C-Means clustering mechanism and enhanced sqrt-cosine similarity.The proposed

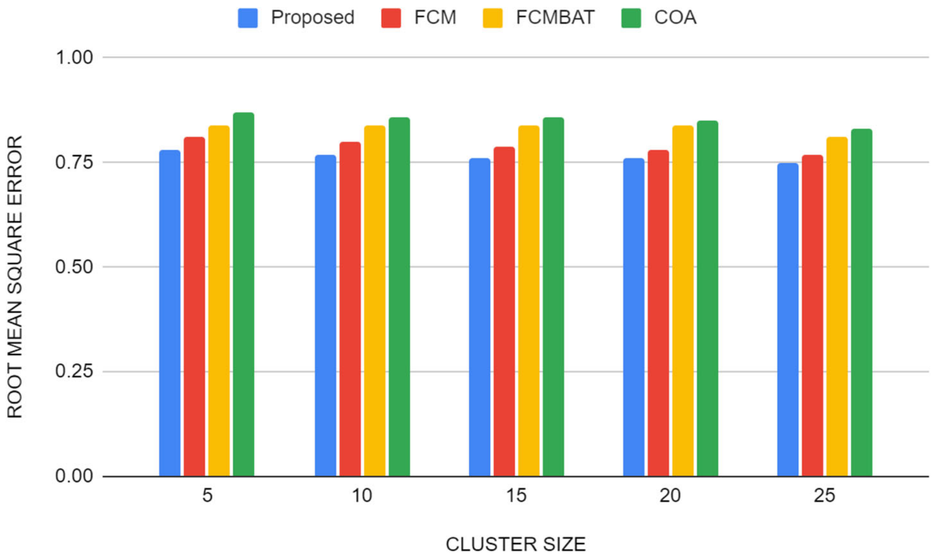


Fig. 3 RMSE comparison analysis

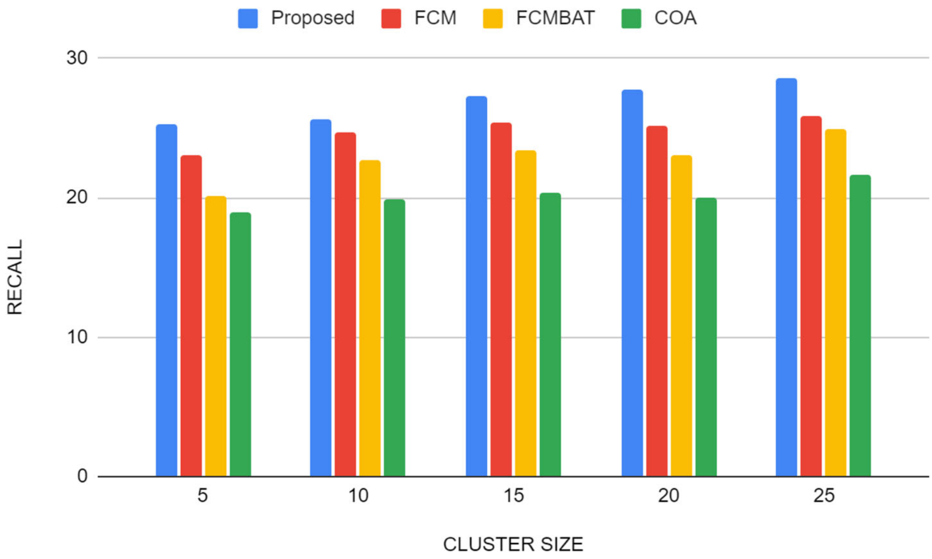


Fig. 4 Recall comparison analysis

scheme emphasized and presented the critical role of the KL divergence-based cluster ensemble factor in improving clustering stability and robustness. For prediction, the enhanced sqrt-cosine similarity was used to calculate successful related neighbor users. The performance of Recommendation is improved when KLD-FCM is combined with improved sqrt-cosine similarity.The proposed KLD-FCM scheme was found to be superior in recommendation Accuracy compared to the COA, FCM-BAT, FCM and ICF approaches, as well as some non-clustering based methods considered for study, when tested on the Movielens dataset in terms of MAE, RMSE, SD, and Recall. With the specified number of clusters, it is capable of

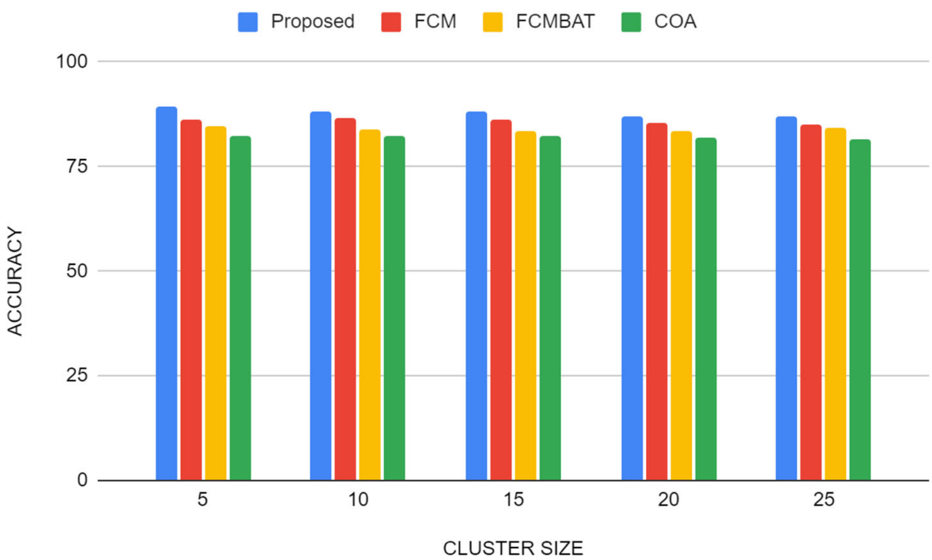


Fig. 5 Accuracy analysis

providing accurate and customized movie recommendation systems. In future work, the proposed design has to be tested with different datasets.

Declarations

Conflict of interest The author(s) propose a clear no conflict of interest involved in this research work in form of publication in this Journal Multimedia Tools and Applications.

References

1. Antony Vijay J, Anwar Basha H, Arun Nehru J (2021) A dynamic approach for detecting the fake news using random forest classifier and NLP. In: In computational methods and data engineering. Springer, Singapore, pp 331–341
2. Asadi E & Charkari N (2012). Video summarization using fuzzy c-means clustering. ICEE 2012 - 20th Iranian Conference on Electrical Engineering. 690–694. https://doi.org/10.1109/IranianCEE.2012.6292442
3. Basha SM, Rajput DS (2019). Survey on evaluating the performance of machine learning algorithms: past contributions and future roadmap. In deep learning and parallel computing environment for bioengineering systems. Academic Press, Cambridge, pp 153–164
4. Clement J. (2020). Impact of recommendation engine on video-sharing platform -YouTube. https://doi.org/10.13140/RG.2.2.15746.50882
5. Cui L & Dong L & Fu X & Wen Z & Lu N & Zhang G. (2016). A video recommendation algorithm based on the combination of video content and social network: CONTENT AND SOCIAL NETWORK BASED VIDEO RECOMMENDATION. Concurrency and Computation: Practice and Experience. 29: https://doi.org/10.1002/cpe.3900.
6. Davidson J, Liebald B, Liu J, Nandy P, Vleet T, Gargi U, Gupta S, He Y, Lambert M, Livingston B, Sampath D (2010). The YouTube video recommendation system. 293–296. https://doi.org/10.1145/1864708.1864770
7. De Vriendt J, Degrande N, Verhoeyen M (2011) Video Content Recommendation: An Overview and Discussion on Technologies and Business Models. Bell Labs Tech J 16:235–250. https://doi.org/10.1002/bltj.20513
8. Deldjoo Y. (2019). Enhancing video recommendation using multimedia content. https://doi.org/10.1007/978-3-030-32094-2_6
9. Deldjoo Y, Elahi M, Quadrana M, Cremonesi P (2015). Toward Building a Content-Based Video Recommendation System Based on Low-Level Features. https://doi.org/10.1007/978-3-319-27729-5
10. Deldjoo Y, Elahi M, Cremonesi P, Garzotto F, Piazzolla P, Quadrana M (2016) Content-Based Video Recommendation System Based on Stylistic Visual Features. J Data Semant 5:1–15. https://doi.org/10.1007/s13740-016-0060-9
11. Deldjoo Y, Schedl M, Cremonesi P, Pasi G (2020) Recommender Systems Leveraging Multimedia Content. Comput Surv 53:1–38. https://doi.org/10.1145/3407190
12. Gupta M, Thakkar A, Gupta V, Rathore DP (2021). Movie Recommender System Using Collaborative Filtering. 978–979
13. Homann L, Martins D, Vossen G, Kraume K (2018) Enhancing traditional recommender systems via social communities. Vietnam J Comput Sci 6. https://doi.org/10.1142/S2196888819500040
14. Huang Y, Cui B, Jiang J, Hong K, Zhang W, Xie Y (2016). Real-time Video Recommendation Exploration. 35–46. https://doi.org/10.1145/2882903.2903743.
15. Kamran M, Shah SS, Baig MN, Khan RH (2020). A movie recommender system by combining both content based and collaborative filtering algorithms
16. Khadse VP, Basha SM, Iyengar N, Caytiles R (2018) Recommendation engine for predicting best rated movies. Int J Adv Sci Technol 110:65–76
17. Lu W & Chung FL (2016). Computational Creativity Based Video Recommendation. 793–796. https://doi.org/10.1145/2911451.2914707.
18. Mercanoglu O & Yıldırım Z (2017). Video Recommendation System Using Collaborative Filtering
19. Mohamed A, Sherif A, Osama F, Roshdy Y, Hassan MA, El Ashmawi WH (2020). A new challenge on video recommendation by content. https://doi.org/10.1109/ICCES48960.2019.9068169.
20. Patil, Lalit. (2016). Fuzzy C means clustering MATLAB code. https://doi.org/10.13140/RG.2.1.3924.9046.

21. Ramezani M, Yaghmaee F (2016) A novel video recommendation system based on efficient retrieval of human actions. Physica A: Statistical Mechanics and its Applications 457. https://doi.org/10.1016/j.physa.2016.03.101
22. Shah P, Sanghvi S (2020) Video Recommender System
23. Tohidi N, Dadkhah C (2020) Improving the performance of video Collaborative Filtering Recommender Systems using Optimization Algorithm. Int J Nonlinear Anal Appl (IJNAA) 11:283–295. https://doi.org/10.22075/IJNAA.2020.19127.2058
24. Zhou X, Chen L, Zhang Y, Cao L, Huang G, Wang C (2015). Online Video Recommendation in Sharing Community. 1645–1656. https://doi.org/10.1145/2723372.2749444.

INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

Intelligent Particle Swarm Optimization Based Resource Provisioning Technique in Cloud Computing

B Ramana Reddy[1,2]*, **M Indiramma**[3], **N Nagarathna**[3]

1 Research Scholar, Department of CSE, B.M.S. College of Engineering, Bangalore, India
2 Assistant Professor, Department of CSE, Chaitanya Bharathi Institute of Technology, Hyderabad, India
3 Professor, Department of CSE, B.M.S. College of Engineering, Bangalore, India

OPEN ACCESS

Received: 12-02-2023

Accepted: 04-04-2023

Published: 28-04-2023

Citation: Reddy BR, Indiramma M, Nagarathna N (2023) Intelligent Particle Swarm Optimization Based Resource Provisioning Technique in Cloud Computing. Indian Journal of Science and Technology 16(16): 1241-1249. https://doi.org/10.17485/IJST/v16i16.308

* **Corresponding author**.

bramanareddy_cse@cbit.ac.in

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Objectives: To find the optimal allocation of resources and minimize the overall cost while meeting the performance requirements of the applications. **Methods:** The proposed Intelligent PSO-based resource optimization in cloud computing evaluates the quality of the solutions based on their resource allocation parameters. Cloud Sim software is used as a simulation tool for testing and evaluating new solutions and strategies in the cloud. The Closest Data Center Service Broker Policy is implemented in Cloud Simulation. **Findings:** The proposed technique assigns workloads effectively on available resources with an improvement of 10.46% in electricity consumption. **Novelty:** The algorithm can further be employed for identifying the unused VM's in the Data Center to reduce the cost.

Keywords: Cloud computing; Resource Scheduling; Load balancing; Particle Swarm Optimization; Quality of Service

1 Introduction

In the field of computer science, cloud computing technology is showing phenomenal growth due to the advancement of the internet[1]. Cloud computing provides infrastructure, platform, and software as services. Cloud resources are providing customers with a pay-as-you-use model. To provide quality services to customers, there is a Service Level Agreement (SLA) between the customers and the cloud service providers as shown in Figure 1. Cloud service providers need to verify if a sufficient number of resources are available to customers to ensure that QoS requirements like execution time, deadline, and budget restrictions are met. However, running too many applications on a single resource can result in a drastic performance loss, which deters cloud consumers. It is challenging to match workloads to the right resources for cloud execution. The existing approaches of Resource optimization in cloud computing involves managing multiple resources and balancing different performance metrics, which can be complex and challenging. Developing an effective optimization algorithm requires significant expertise and resources. For effective resource use, three primary QoS limitations must be taken into account[2].

1. Must meet deadlines with minimum execution time and cost
2. Minimum electricity consumption
3. User satisfaction

A PSO-based resource scheduling algorithm, schedules workloads in a cloud environment to save execution time, cost, and energy. By figuring out how workloads and resources interact in a cloud setting was expanded. Experimental findings enhance QoS metrics like resource availability, reliability, latency, and resource consumption to the maximum extent possible [3].

The Particle Swarm Optimization and Genetic Algorithm (PSO-GA) was used to make runtime decisions for exploring the objectives of the resource allocation plan [4]. The issue of resource allocation for cloud-based services is addressed by introducing the workload time window. Further, it builds a calculation model for optimizing VM resource allocation plans [5]. The PSO-GA-based method observed the advantages of PSO and GA and improved their inadequacies in population diversity, search range, and convergence speed. The algorithm strategy objective takes the current and future workloads into the process of producing resource allocation plans.

In the proposed research work, the intelligent Particle Swarm Optimization focuses on minimizing execution time, cost and electricity consumption.

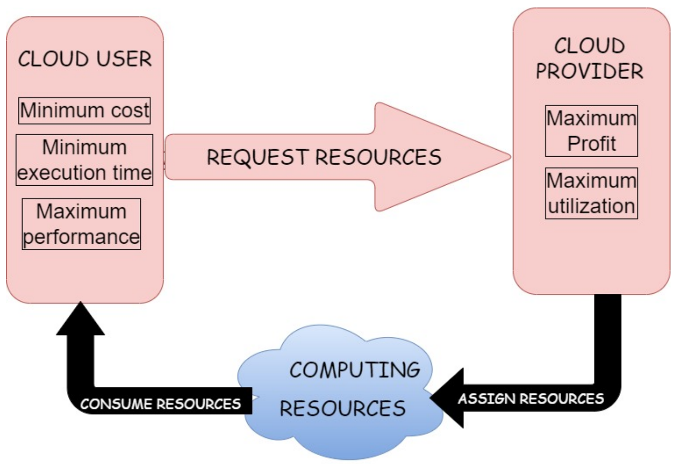


Fig 1. Resource Request in Cloud Computing

2 Methodology

2.1 Proposed Intelligent Particle Swarm Optimization-based resource

2.1.1 Scheduling Technique

In cloud computing, resource scheduling is important for resource management [6]. It directs the allocation of cloud workloads to cloud resources. After, applications are scheduled to cloud resources using an Intelligent Particle Swarm optimization-based heuristic framework, which lowers the cost of computation and data transfer. Most existing research takes into account of cloud computing's fundamental characteristics in order to execute heterogeneous cloud work loads quickly and affordably [7]. The population in the I-PSO algorithm is defined as particles, where the particles are initialized randomly. In every new generation, the fitness value of each particle is calculated and the two values of the particles need to be calculated: 1. Local best(L_{best}), 2. Global best(G_{best}). Where the L_{best} of a particle is the best result reached by the particle so far, and the G_{best} is the best result among the whole population. Intelligent PSO optimization technique works on global search [8]. Every single particle controls its own independent course according to the local and global best in every generation. Numerous NP-hard issues, including task distribution and resource scheduling, will be resolved by intelligent PSO.

Workloads : Workloads that are initiated by the user are placed in a queue for execution and processing purposes.

Resource Manager: The resource manager maintains information about resources, QoS, and SLA.

Quality of Service: Quality of Services like availability, latency, resource utilization, and reliability must be maintained.

Service Level Agreement : Provides information about suitable service level agreements between customers and cloud service providers.

Workload Analyzer: Workload Analyzer identifies different characteristics of cloud workloads.

Resource details: Maintains resource details such as availability of virtual machines, size of virtual memory, cost of cloud resources, and type of cloud resource.

Resource provisioner: It provides workload to the demanded resources for their execution in the cloud only if the resources are available in the resource pool.

Resource scheduler: It executes all the workloads on a provisioned resource efficiently.

The Proposed architecture of resource optimization is described in Figure 2.

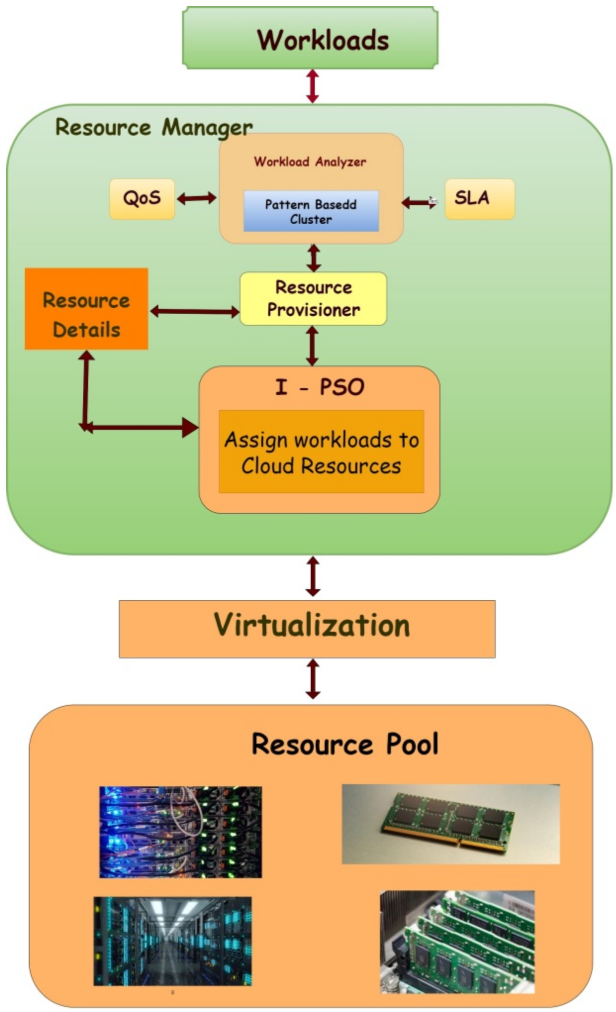


Fig 2. Architecture of proposed Resource optimization

2.1.2 Objective Function

The fitness value is calculated based on execution time, cost, and electricity consumption as shown in equation 1.

$$\text{fitness } = c1 * \text{ execution time } + c2 * \text{ execution cost } + c3 * \text{ electricitycon} \tag{1}$$

$0 \leq \text{c1} \leq 1, 0 \leq \text{c2} \leq 1$ and $0 \leq \text{c3} \leq 1$

where c1, c2, and c3 are weights to prioritize the components of the fitness function.

Execution time is the time taken to execute the workload on the allocated resources as follows:

$$\text{Execution }_{\text{time}} = \min\left(t\left(w_i, r_i\right)\right) \text{ for } 1 \leq i \leq n \tag{2}$$

where $t(w_i, r_i)$ is the time taken by workload w_i executed by resource r_i.

Execution cost is the cost of the workloads executed by the assigned resources in the following equation.

$$\text{Execution}_{\text{ cost}} = \min\left(c\left(w_i, r_i\right)\right) \text{ for } 1 \leq i \leq n \tag{3}$$

where $c(w_i, r_i)$ is the cost of workload w_i executed by resource r_i. electricity consumption is the electricity consumed in resource utilization calculated by

$$\text{electricity}_{\ co\ n} = \text{electricity}_{\ vm} + \text{electricity}_{\text{ memory}} + \text{electricity}_{\text{ misc}} \tag{4}$$

Where electricity_{vm} is the virtual machine electricity consumption, $\text{electricity}_{memory}$ is the electricity consumption for memory operations and $\text{electricity}_{misc}$ is the electricity consumption for fans and other miscellaneous parts.

$$\text{electricity}_{\ vm} = vm_{\text{idle}} + \left(vm_{\text{running}} - vm_{\text{idle}}\right) \times vm_{\text{processor}} \tag{5}$$

where vm_{idle} denotes the idle state of the virtual machine, $vm_{running}$ denotes the virtual machine running time power consumption, and $vm_{processor}$ denotes the virtual machine processor capacity.

2.2 Proposed Algorithm

In this section, the description regarding I-PSO based algorithm for resource provisioning technique in the cloud environment is shown in Algorithm 1.

Algorithm 1: Proposed Intelligent PSO-based resource scheduling Algorithm

Result: intelligent mapping of workloads to the VMs.

```
initialize VMs
initialize workloads
initialize a random feasible solution
for i=1 to PS
do
Pv ← RV
Pp ← RP
if fitness(G_best) > f itness(P )
then
G_best ← P
end
end
while iterations are not reached max
do
for P ∈Ppop
do
Pv ← update velocity
Pp ← update position
if f itness(P ) < f itness(L_best)
then
L_best ← P
end
if f itness(Lbest) < f itness(Gbest)
then
Gbest ← Lbest
end
end
return(Gbest)
end
if any unassigned VMs then
VMs moved to sleep mode.
```

32 end
Intelligent PSO Terminology:
The notations used in an I-PSO algorithm are presented in Table 1.

- **Particle:** A particle is similar to a flock of birds searching for food. Every particle has velocity. Fitness values calculate a particle's performance. For the proposed I-PSO algorithm, cloud workloads are considered particles.
- **Population size:** In the proposed I-PSO, available resources in the cloud are considered as population size.
- **Random velocity:** Every particle's velocity is updated with L_{best} and G_{best} values.
- **Particle velocity:**The particle's velocity is calculated based on particle position.
- **Particle position:** positions depend on the submission status, waiting state, ready state, execution state, and completion state.
- **Global best (G_{best}):** Best position of a particle among the whole group of particles.
- **Local best (L_{best}):** Best position reached by a particle

Table 1. Notations and its description

Notation	Description
S	Number of particles
PS	Population size
RV	Random velocity
RP	Random position
Pv	Particle velocity
Pp	Particle position
Ppop	Particle population
G_{best}	Global best position
L_{best}	Local best position

In Figure 3, all the algorithm steps are described in the form of a flow chart.

3 Results and Discussion

The experiment has been conducted in a simulator with size of 700MB image conversion cloud workload. Once the simulation starts a sequence of steps will be execute as per the mentioned order.

1. Cloud user submits workload details like name and type of workload. Workload analyzer will submit these details to Resource Manager.
2. Resource manager processes workload details and then asks the user for budget restrictions and deadline details.
3. Once the cloud user submits budget and deadline restrictions the Resource manager will generate tentative schedule and cost.
4. Cloud user sends a confirmation of the details, SLA agreements to the Resource Manager [9].
5. Cloud user pays requested amount and executes the workloads on allocated resources

The performance of the proposed I-PSO resource provisioning technique is compared with the existing scheduling algorithms [10,11].

Results – Execution time: In Table 2 results compared at 45 workloads, execution time in I-PSO is 2.74% lesser than ACO, 4.31% lesser than GA. At 90 workloads, execution time in I-PSO is 3.53% lesser than ACO, 5% lesser than GA. Figure 4 shows that execution time of I-PSO is better than ACO and GA.

Results – Execution cost: In Table 3 results compared at 45 workloads, the cost incurred in I-PSO is $260, ACO is $266 and GA is $292. At 90 workloads the cost incurred in I-PSO is $412, ACO is $422 and GA is $442. Figure 5 shows that the execution cost of I-PSO is minimum compared to the ACO and GA.

Results – Electricity consumption: In Table 4 results compared at 45 workloads the I-PSO consumes 8.97% lesser than ACO, 14.10% lesser than GA. At 90 workloads the I-PSO consumes 10.46% lesser than ACO, 16.27% lesser than GA. Figure 6 shows that the electricity consumption is minimum compared to the ACO and GA.

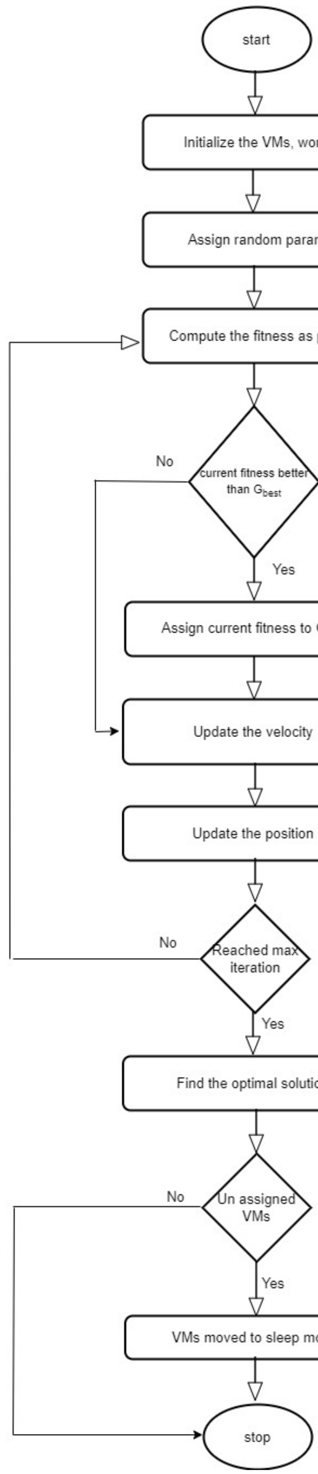


Fig 3. Flowchart of the proposed system

Table 2. Comparison of I-PSO Execution time (ms) with GA and ACO works

No. of work loads	Execution Time (ms)		
	GA	ACO	I-PSO
15	225	221	198
30	410	402	380
45	532	524	510
60	698	788	765
75	950	948	918
90	1218	1201	1160

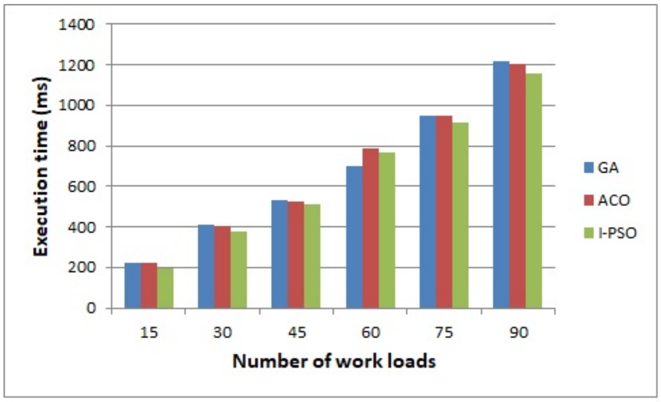


Fig 4. Execution time vs no.of work loads

Table 3. Comparison of I-PSO Execution cost($) with GA and ACO works

No. of work loads	Execution cost ($)		
	GA	ACO	I-PSO
15	196	176	169
30	230	204	198
45	292	266	260
60	332	307	297
75	368	346	336
90	442	422	412

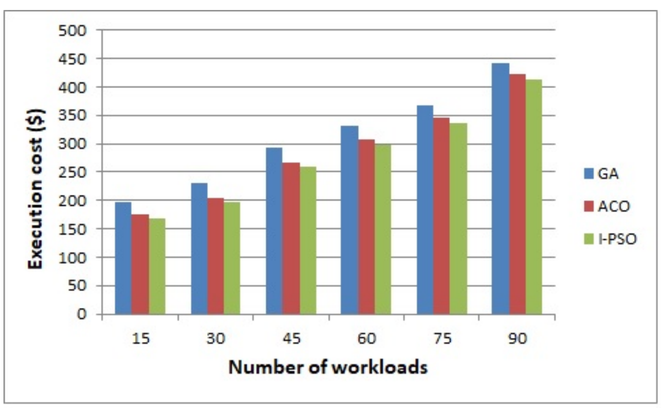


Fig 5. Execution cost vs no.of work loads

Table 4. Comparison of I-PSO electricity consumption (kwh) with GA and ACO works

No. of work loads	Electricity Consumption (Kwh)		
	GA	ACO	I-PSO
15	82	76	70
30	88	82	76
45	89	85	78
60	92	87	81
75	96	92	84
90	100	95	86

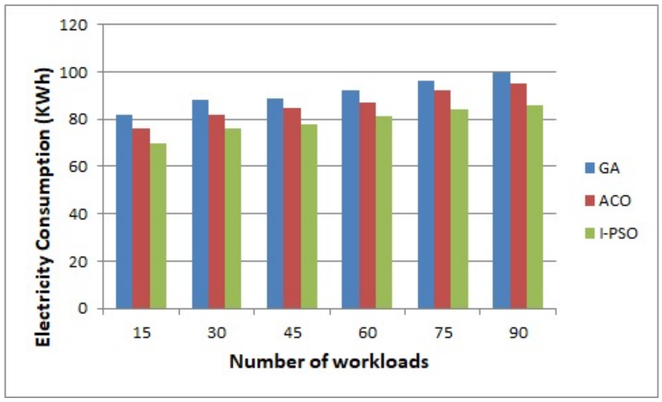


Fig 6. Electricity consumption vs no.of workloads

4 Conclusion

In this research work, different techniques such as load balancing and resource allocation methods are provided which ensures that resources are used optimally. Furthermore, the Proposed Intelligent PSO algorithm can help predict idle virtual machines, enabling providers to proactively allocate resources. However, the I-PSO algorithm is unable to predict the upcoming workloads due to ongoing changes of user requirements.

To further improve this research work, the computation of the fitness value can be more accurate by considering the priority of virtual machines. Thus, the highest priority virtual machine will be assigned first and this result in a lower execution time. The future scope will require a continuous improvement and innovation to address emerging challenges and technologies. One key area of focus will be the development of several sophisticated machine learning algorithms that can accurately predict resource demands and dynamically allocate resources in real-time. This will enable cloud providers to achieve higher levels of efficiency, reduce costs, and improve overall performance.

References

1) Mishra SK, Sahoo B, Parida PP. Load balancing in cloud computing: A big picture. *Journal of King Saud University - Computer and Information Sciences*. 2020;32(2):149–158. Available from: https://doi.org/10.1016/j.jksuci.2018.01.003.
2) Singh A, Kumar R. Performance Evaluation of Load Balancing Algorithms Using Cloud Analyst. *2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*. 2020;p. 156–162. Available from: https://doi.org/10.1109/Confluence47617.2020.9058017.
3) Kruekaew B, Kimpan W. Multi-Objective Task Scheduling Optimization for Load Balancing in Cloud Computing Environment Using Hybrid Artificial Bee Colony Algorithm With Reinforcement Learning. *IEEE Access*. 2022;10:17803–17818. Available from: https://doi.org/10.1109/ACCESS.2022.3149955.
4) Chen Z, Yang L, Huang Y, Chen X, Zheng X, Rong C. PSO-GA-Based Resource Allocation Strategy for Cloud-Based Software Services With Workload-Time Windows. *IEEE Access*. 2020;8:151500–151510. Available from: https://doi.org/10.1109/ACCESS.2020.3017643.
5) Reddy RB, Indiramma M. Efficient Throttled load balancing algorithm to improve the response time and processing time in data center. *Concurrency And Computation-Practice & Experience*. 2022;34(23). Available from: https://doi.org/10.1002/cpe.7208.
6) Malik N, Sardaraz M, Tahir M, Shah B, Ali G, Moreira F. Energy-Efficient Load Balancing Algorithm for Workflow Scheduling in Cloud Data Centers Using Queuing and Thresholds. *Applied Sciences*. 2021;11(13):5849. Available from: https://doi.org/10.3390/app11135849.
7) Sweekriti M, Shetty S, Shetty. Analysis of load balancing in cloud data centers. *Journal of Ambient Intelligence and Humanized Computing*. 2019;p. 1–9. Available from: https://doi.org/10.1007/s12652-018-1106-7.

8) Kothilaxman RR, Lathigara A, Aluvalu R, Viswanadhula UM. PGWO-AVS-RDA: An intelligent optimization and clustering based load balancing model in cloud. 2022. Available from: https://doi.org/10.1002/cpe.7136.
9) Samriya JK, Patel SC, Khurana M, Tiwari PK, Cheikhrouhou OK. Intelligent SLA-Aware VM Allocation and Energy Minimization Approach with EPO Algorithm for Cloud Computing Environment. *Mathematical Problems in Engineering*. 2021;2021:1–13. Available from: https://doi.org/10.1155/2021/9949995.
10) Singh H, Tyagi S, Kumar P, Gill SS, Buyya R. Metaheuristics for scheduling of heterogeneous tasks in cloud computing environments: Analysis, performance evaluation, and future directions. *Simulation Modelling Practice and Theory*. 2021;111:102353. Available from: https://doi.org/10.1016/j.simpat.2021.102353.
11) Ragmani A, Elomri A, Abghour N, Moussaid K, Rida M. An improved Hybrid Fuzzy-Ant Colony Algorithm Applied to Load Balancing in Cloud Computing Environment. *Procedia Computer Science*. 2019;151:519–526. Available from: https://doi.org/10.1016/j.procs.2019.04.070.

Sustainable Cities and Society 97 (2023) 104722

Contents lists available at ScienceDirect

Sustainable Cities and Society

journal homepage: www.elsevier.com/locate/scs

An Optimized Bio-inspired Localization Routing Technique for Sustainable IIoT Networks & Green Cities

Premkumar Chithaluru [a], Fadi Al-Turjman [b,c], Thompson Stephan [d,*], Manoj Kumar [e,f], Sunil Kumar [g]

[a] Department of Computer Science and Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad 500075, India
[b] Research Center for AI and IoT, Faculty of Engineering, University of Kyrenia, Mersin 10, Turkey
[c] Artificial Intelligence Engineering Dept., AI and Robotics Institute, Near East University, Mersin 10, Turkey
[d] Department of Computer Science and Engineering, Faculty of Engineering and Technology, M. S. Ramaiah University of Applied Sciences, Bangalore, Karnataka 560054, India
[e] School of Computer Science, FEIS, University of Wollongong in Dubai, Dubai Knowledge Park, Dubai, United Arab Emirates
[f] MEU Research Unit, Middle East University, Amman, 11831, Jordan
[g] Department of Cybernetics, School of Computer Science, University of Petroleum & Energy Studies (UPES), Bidoli, Dehradun 248007, India

ARTICLE INFO

Keywords:
IIoT
Sustainable
ICSO
PSO
Localization
Bio-inspired
Green cities

ABSTRACT

The industrial Internet of Things (IIoTs) network life is shortened due to sensor node (SN) energy limitations and computational capability. As a result, optimum node location estimation and efficient energy usage are two critical IIoT requirements. This work reduces energy consumption by performing node localization and cluster-based routing using an improved evolutionary algorithm called Cat Swarm Optimization (CSO). First, the CSO method is used to optimize the bio-inspired node's location. Second, to conserve SN energy in the IIoT network, a cluster-based routing technique is used. The objective function is defined as minimizing the average distance between the cluster and its SNs while selecting the most energy-efficient Cluster Head (CH). In terms of fitness value, the Improved CSO (ICSO) algorithm outperforms the Particle Swarm Optimization (PSO) algorithm. In this paper, real-time test-bed analysis was used to investigate the performance of both node localization and energy-efficient clustering. When it comes to achieving sustainable IIoT and green cities, the findings show that ICSO outperforms in terms of convergence rate and network lifetime.

1. Introduction

The development of smart devices is facilitated by the creation of a user-friendly environment. These devices should be self-aware and communicative. The Internet of Things (IoT) connects intelligent devices to perform a task by utilizing smart networking protocols (Wu et al., 2023). The rapid advancement of wireless communication technologies has piqued the interest of both industry and academia in real-world solution research and development. Furthermore, IoT is a powerful infrastructure-less networking unit comprised of low-power SNs organized in an ad-hoc system (Tanwar, Balamurugan, Saini, Bharti, & Chithaluru, 2022). These SNs can collect and process data from their intended environment to communicate with one another. Enemy battlefield monitoring, ecological parameter monitoring, off-shore energy sector monitoring, biological detection, industrial diagnostics, and other sensors are randomly deployed in a physical environment that serves as the sensing layer for IIoT applications (Shit, Sharma, Puthal, & Zomaya, 2018).

The IIoT is a network of independent SNs distributed in a nested application that can sense, determine, and update data in the Base Station (BS) (Chithaluru, Al-Turjman, Kumar, & Stephan, 2021). Fig. 1 depicts a typical IIoT application in various emerging technologies. Each SN in the IIoT is made up of a sensing unit for sensing environmental parameters, an antenna for transmission and reception, and a processor and memory device (Lashkari, Rezazadeh, Farahbakhsh, & Sandrasegaran, 2018). Nonetheless, the IIoT components are designed with more constraints, particularly in terms of energy and processing capabilities. The main limitation of IIoT is the sensor unit's power supply. When deployed in a hostile environment, the batteries that power the SNs are not easily replaceable. As a result, energy consumption is a difficult subject that necessitates extensive research to extend the IIoT's lifespan (Chithaluru, Kumar, Singh, Benslimane, & Jang, 2021).

* Corresponding author.
E-mail addresses: bharathkumar30@gmail.com (P. Chithaluru), fadi.alturjman@neu.edu.tr (F. Al-Turjman), thompsoncse@gmail.com (T. Stephan), wss.manojkumar@gmail.com (M. Kumar), drskumar.cs@gmail.com (S. Kumar).

https://doi.org/10.1016/j.scs.2023.104722
Received 31 August 2022; Received in revised form 10 April 2023; Accepted 7 June 2023
Available online 4 July 2023

Fig. 1. Green & sustainable IIoT applications.

SNs deployed to collect data and transmit it to an assigned BS in a self-organized manner, despite their energy constraints and limited computing, storage, and communication link capacities (Joshi et al., 2023). This intense resource limitation requires mechanisms to consume low power during the gathering and routing of the data. The energy reduction goal in IIoT allows the sensors to communicate with the BS through a multi-hop routing technique. Nevertheless, till today, it is been a challenge to have a low-power system to sense and route the data to the BS in an effective manner. Hence, the design and maintenance of sensor networks required a scalable architecture and powerful organizational techniques (Chithaluru, Al-Turjman, Stephan, Kumar, & Mostarda, 2021).

1.1. Localization in IIoT

The estimation of the location of deployed nodes in IIoTs is still an important part of current research. As shown in Fig. 2, the localization technique in IIoT allows for estimating the precise coordinate position of SNs deployed in a field of interest. IIoTs are used for a variety of monitoring tasks in a variety of environments, including target tracking on enemy battlefields and rescue and disaster operations. The location awareness of the SNs is important to the application's localization techniques in this monitoring and can be classified based on the design and application plans such as range information, connectivity, anchor information, and computational method.

Existing localization techniques are divided into two categories: distance estimation and combining and angle estimation and combining. Referring to Table 1 for acronyms and variables.

1.1.1. Distance/angle evaluation

The following schemes are used to calculate the distance or angle between SNs.

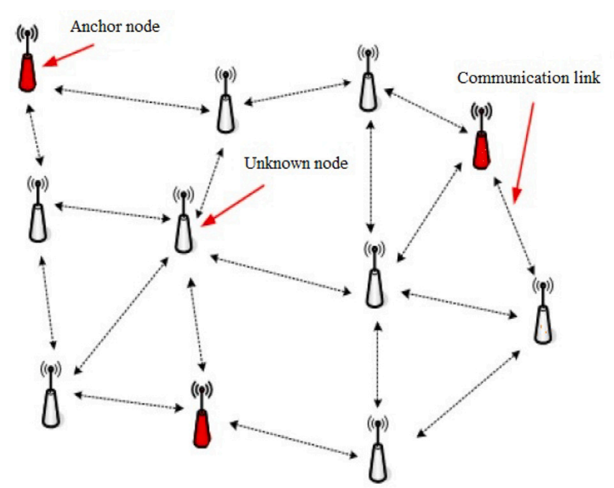


Fig. 2. Localization in green IIoT network.

- **Received Signal Strength Indicator (RSSI):** The propagation loss is calculated using the received signal power. To interpret the propagation loss in terms of distance, hypothetical models are used. This method is widely used for Radio Frequency (RF) signals. The signal strength α_{Rec} is calculated using Friis' expression, as shown in Eq. (1).

$$\alpha_{Rec} = \alpha_{Tran} \times \Gamma_{Tran} \times \Gamma_{Rec} \times (\frac{\Lambda}{4\pi D})^2 \tag{1}$$

It is clear from Eq. (1) that the α_{Rec} limits with distance. RSSI is used by the SNs to translate the RSS. Finally, the RSSI is

Table 1
Acronyms.

Symbol	Meaning
α_{Rec}	Signal Strength
α_{Ref}	Reference power
$\delta\tau$	Time difference of transmittedand received signals
ε_R	Velocity of RF
D	Distance
CH_{Prob}	Probability of CH
η	Number of SNs
c	Number of round
ς_{Prob}	Initial % of CHs
ϵ_{res}	SN's remaining energy
ϵ_{max}	Maximum energy of SN's
$\epsilon_i(x)$	Primary energy of SNs
ρ	Cluster span
ψ	Hop count

calculated using Eq. (2).

$$RSSI = 10 \times log(\frac{\alpha_{Rec}}{\alpha_{Ref}}) \tag{2}$$

where α_{Ref} is the reference power, which is typically 1 mW. The RSSI is measured in decibels (dBm). The RSSI is an important factor in range measurements because it can be calculated simply by analyzing the received signal power. However, the RSSI method's measurement accuracy is particularly sensitive to multi-path fading, which can cause huge errors in subsequent computation techniques and result in ineffective information.

- **Angle of Arrival (AoA):** The nodes in the AoA method approximate the reception angle of the signal and use simple arithmetic alterations to determine node positions. AoA can be solved using the triangulation method. The AoA method is susceptible to measurement noise, and it also requires awareness of the reference direction. As a result, the AoA method necessitates antenna arrays for each node, making the system costly.
- **Time-Difference-of-Arrival (TDoA):** The TDoA technique calculates distance using propagation time and a known signal propagation frequency. The TDoA scheme can be applied to a wide range of signals, from acoustic to RF. The variable frequency signal is an example of RF and acoustic signals transmitted concurrently with known time and received at times τ_1 and τ_2, respectively. The time difference between transmitted and received signals is $\delta\tau$, and the velocity of the RF/acoustic signal is ϵ_ρ and ϵ_S, respectively, and the Eq. (3) is fulfilled.

$$d_{\tau-\rho} = \frac{\delta\tau \times \epsilon_S}{1 - \frac{\epsilon_S}{\epsilon_\rho}} \tag{3}$$

As $\epsilon_S \ll \epsilon_\rho$, then Eq. (4) is computes,

$$d_{T\tau-\rho} = \delta\tau \times \epsilon_S; \tag{4}$$

1.1.2. Distance and angle combining

The information obtained by using the methods described above aids in determining the coordinates of the nodes in phase two. The key methods used in phase two of localization are discussed further below.

- **Trilateration:** This technique makes use of the measured distances between three or more nodes, as shown in Fig. 3.
The basic sphere formula is given in the following set of equations (Eqs. (5)–(8)), which assist in determining the position estimate.

$$D^2 = p^2 + q^2 + r^2 \tag{5}$$

For sphere 1

$$D_1^2 = (p - p_1)^2 + (q - q_1)^2 \tag{6}$$

For sphere 2

$$D_2^2 = (p - p_2)^2 + (q - q_2)^2 \tag{7}$$

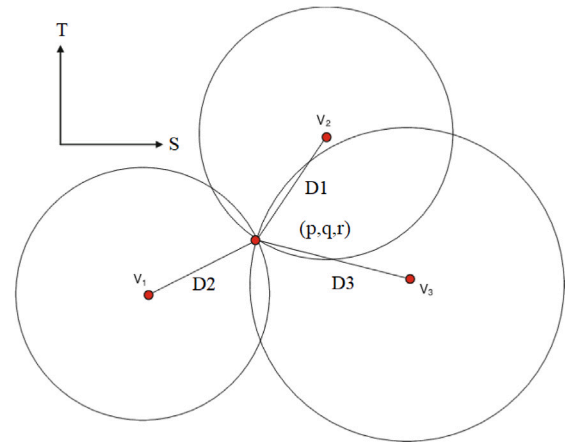


Fig. 3. Working principle of trilateration.

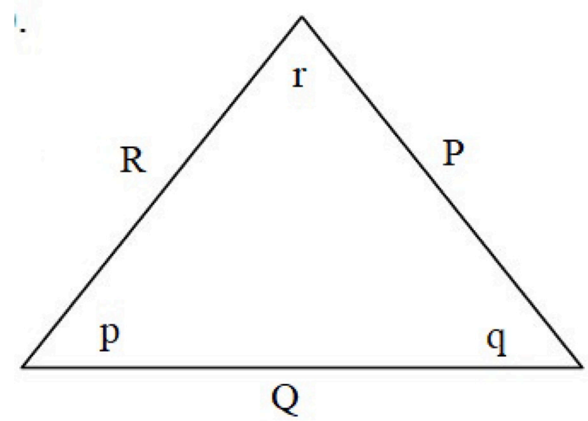


Fig. 4. Triangulation.

$$D_3^2 = (p - p_3)^2 + (q - q_3)^2 \tag{8}$$

- **Triangulation:** To calculate the coordinate positions, this scheme employs two angles of measurement and one distance estimate. The angle is measured using triangulation, as shown in Fig. 4. The geometry of beacon nodes and ordinary nodes participating in localization is used to estimate the accuracy of the triangulation method. Eqs. (9) and (10) represent the corresponding position estimations.

$$\frac{p}{sinP} = \frac{q}{sinQ} = \frac{r}{sinR} \tag{9}$$

$$\begin{aligned} r^2 &= p^2 + q^2 - 2pqCosR \\ q^2 &= p^2 + r^2 - 2prCosQ \\ p^2 &= q^2 + r^2 - 2qrCosP \end{aligned} \tag{10}$$

1.2. Motivation

IIoTs are distinct from other wireless networks due to their unique nature, various limitations, and high-level applications, posing numerous research challenges. Sensor deployment, network coverage, connectivity, and energy resources are the primary factors that determine an IIoT's efficiency. Localization is a critical process in the IIoT for reporting the origin of measures and routes. It also aids in network coverage and sensor querying. The RSSI method is used in localization to estimate the distance of an unknown SN and can be improved by using a bio-inspired algorithm, as highlighted. The SN uses energy to perform tasks such as measuring, processing information, transmitting, and receiving. The data transmission phases, in particular, consume a significant amount of energy. Because the SN is a low-power microelectronic device, its lifespan is solely determined by the life of the battery. Routing protocols play a critical role in energy minimization in most IIoTs, and they decide on the path of information flow from source to destination. To save energy, the data transfer method must be more efficient. The data aggregation method also saves energy in the sensor network by compressing data before sending it to the BS, which reduces the amount of information transmitted.

The clusters formed in the peer existing cluster-based routing protocols are not uniform; some clusters have more nodes than others. In CH with large clusters, the disparity of the nodes formed within the cluster group causes more imbalanced traffic and a rapid reduction in energy. As a result, to improve the proposed protocol, new optimization techniques must be developed to obtain the global optimum of network establishment. In this paper, simulation, and experiments are used to investigate IIoT localization and cluster-based routing protocols using optimization algorithms based on bio-inspired techniques, specifically the PSO and ICSO algorithms.

1.3. Problem statement

Because IIoT is highly dynamic and sensitive to energy consumption, clustering becomes a difficult task. When designing the operation of an IIoT to gain an extended lifetime, energy savings are critical. Thus, awareness of the deployed SNs' location, residual energy, and power consumption of various SN components is extremely important. It is critical to research and selects an efficient algorithm to help organize the operation of the IIoT in terms of energy savings. These elements are taken into account and investigated.

The main goal is to perform effective node localization and to extend the IIoT lifetime by utilizing a novel energy management technique. Bio-inspired algorithm models are appropriate for the IIoT's dynamic nature, which can be achieved by incorporating PSO into the ICSO algorithm for both localization and clustering. The swarm-based intelligence techniques used in this work solve the problem effectively. Using an enhanced bio-inspired algorithm, this paper designs, analyses, and tests an optimized node localization and energy-efficient cluster-based protocol for IIoT.

The major contributions of the paper are as follows,

- We proposed a new node localization technique for long-term IIoT networks based on RSSI-distance conversion and optimization with the ICSO technique.
- A simulated cluster-based protocol for sustainable IIoTs is proposed with an emphasis on energy efficiency and network lifetime extension.
- A simulated cluster-based protocol for sustainable IIoTs is suggested with an emphasis on energy efficiency and network lifetime extension.
- To design and implement an ICSO-based centralized cluster-based protocol for sustainable IIoT cities routing.
- The ICSO algorithm is used to solve a multi-objective problem by optimizing cluster formation and CH selection at the same time.
- The results of the ICSO protocol are analyzed and compared to peer existing routing protocols such as RSSI (Mohar, Goyal, & Kaur, 2022) and PSO (Chithaluru, Tiwari, & Kumar, 2021b).

2. Literature review

The focus of this section is on IIoT node localization and the cluster-based routing protocol to provide an energy-efficient IIoT. As a result, the literature studies have focused primarily on the peer existing and recent approaches used in the research area for both localization and clustering. The literature review is divided into two sections: anchor-based techniques and anchor-free techniques. These particulars will be addressed in the following sections:

2.0.1. Anchor based localization

In experimental tests, the cricket application was localized using node arrays, as described in Cheng, Li, and Xu (2022). Chithaluru, Khan, Kumar, and Stephan (2021) developed a lower bound for network localization based on expected error properties, which was then compared to the actual error using multi-alternation. Chithaluru, Singh, et al. (2023) used a semi-definite program to realize node localization based on angles between SNs and connectivity. In the localization technique using AoA presented in Chithaluru, Tiwari, and Kumar (2021a), a well-known Adhoc-based position estimation system is proposed. It is emphasized that a novel localization technique based on AoA information between adjacent nodes is used Sridhar, Chithaluru, Kumar, Cheikhrouhou, and Hamam (2023). This method makes use of multi-hop anchor information. Ou et al. (2022) proposed a phased algorithm for proximity distance mapping and multidimensional scaling. Yan, Yang, Wang, and Shen (2022) used the M-C sampling scheme with a single anchor to interpret different types of AoA, connectivity, and ranging. Luomala and Hakala (2022) used constrained simulated annealing with tunneling transformation to solve a non-tractable posterior distribution.

2.0.2. Anchor free localization

Sabale and Mini (2022) and Niculescu, Palossi, Magno, and Benini (2022) propose a novel anchor-free technique for IIoT applications and analyze a self-positioning method. Both techniques use the SN distances to build a comparative coordinate scheme. In the same year, Arya (2022) described the ABC algorithm, which determines node position one at a time. Jia, Qi, Liu, and Zhou (2022) emphasized an anchor-free location-finding technique that used a corrected signal to measure the AoA and TDoA of SNs. The nodes determine their location based on the estimated location, which increases the probability of observation. Liu, Luo, Wei, and Liu (2022) improved the R-factor accuracy of IIoT localization. Yadav et al. (2023) proposed a cluster-based method for constructing the local coordinate system. Abdulzahra, Al-Qurabat, and Abdulzahra (2023) described a new scheme for improving localization accuracy by using copula cross-correlation and shadow fading noise.

2.1. Performance metrics in localization

2.1.1. Accuracy

The accuracy of localization estimates, which are based on the Euclidean distance between the beacon and the actual position of the unknown SN, is the most important factor. Because of their low measurement error, range-based methods are much more popular in general. The measuring method is used in range-based techniques to calculate the Euclidean distance between nodes. Chithaluru, Al-Turjman, Kumar, and Stephan (2023) developed a range-free localization method that used the shortest path between nodes to estimate where the shortest paths are similar to straight lines, which is not frequent in real life.

In recent years, hybrid localization approaches based on range-based or range-free methods have gained popularity. It is more precise than the standard localization method. Typically, measurement noise is caused by a variety of inputs. As a result, the accuracy of the position estimate for various measurements should be partially self-determining. This sovereignty with multiple measurement types could be undermined by data-fusion schemes such as those proposed by Pradhan, Das, Mishra, and Chithaluru (2023) to create position estimators that are more accurate than a single measurement type.

Hybridized range-free techniques (Chithaluru, Fadi, Kumar, & Stephan, 2023) outperform other fully range-free techniques in terms of estimation accuracy. According to Chen et al. (2020) and Joshi, Albahar, et al. (2022), the hybridized range-based technique outperforms the proposed fully-range-based technique Chithaluru, Jena, Singh, and Ravi Teja (2022). According to Yadav, Chithaluru, Singh, Joshi, et al. (2022), most schemes' accuracy suffers in an obstructed environment because impediments obstruct the SN's line of sight. For the range-based schemes, the emphasis is primarily on physical factors. Localization is more sensitive to environmental factors, as demonstrated by an example of the scheme proposed by Joshi, Chithaluru, et al. (2022).

The deployment strategy of the SNs has a significant impact on localization accuracy. When compared to a randomly distributed sensor network, the uniformly distributed IIoT provides higher accuracy. In anchor-based schemes, the density and placement of beacon nodes also affect localization accuracy. Recent research shows that deploying beacon nodes in a circle around the network improves localization accuracy. Excess anchors can also be installed in the center of the IIoT.

According to the literature, some techniques perform better in a network with higher beacon density, whereas others require only a few beacons. Chithaluru, Stephan, Kumar, and Nayyar (2022) proposed a scheme that works better with a large number of beacon nodes and low communication and computation costs. The majority of localization techniques are sensitive. The number of localizing nodes per unit area (node density) for hop count-based algorithms such as Distance Vector-Hop with correction (CDV) Hop, proposed by Chithaluru, Al-Turjman, et al. (2020), necessitates a higher node density to aid in accurate hop count approximation for distances. Due to connectivity failures and longer distances between SNs, low-density nodes cannot be used in range measurement schemes. Furthermore, when developing a localization technique, node density is an important factor to consider. Shah et al. (2018), on the other hand, proposed a scheme that helps overcome most of the typical problems in IIoT, such as obstacles, low node density, network connectivity, and fault tolerance.

2.1.2. Overhead

Many factors contribute to the overhead of localization techniques, including the number of anchor nodes, computation as well as communication costs, usage of energy, processing speed, required hardware for each node, and many more.

- **Communication cost:** The efficiency of IIoT energy is critical because it determines the network's lifespan. For data transmission and reception, SN requires the most energy. Jain et al. (2022) a hop count-based strategy that necessitates collaboration among adjacent nodes. This method has a high communication cost. As a result, the localization technique should reduce communication costs.
- **Computation cost:** The computation complexity of hybrid techniques grows in proportion to the range and amount of data used in localization. Furthermore, as Joshi, Chithaluru, Singh, et al. (2022) emphasize, most hybrid schemes necessitate higher computation costs than traditional schemes.
- **Hardware cost:** The measurement equipment, node density, and anchor density are all included in the hardware cost. In general, equipment costs more, but the expected accuracy is higher. Of course, range-based techniques necessitate more hardware than range-free techniques; distance measurement between sensors is required (Chithaluru, Jena, Singh, & Nayak, 2022).

2.1.3. Scalability

Another critical factor in validating the localization technique is scalability. When the network or node deployment area is at its largest, it ensures that the location estimation is correct. In a dense network, the wireless signal may become congested, necessitating the use of a complex communication structure. All measurements at a BS are aggregated for processing using centralized techniques. Furthermore, due to the funneling effect, SNs located near the BS will lose their battery source faster than other nodes. Thus, the centralized approaches that must be considered have an impact on scalability. Similarly, as emphasized in Chithaluru (2020), the distributed technique applies to large-scale sensor networks. Chithaluru, Jena, Patra, and Panda (2022) and Jena, Ammoun, and Chithaluru (2022) used the PSO algorithm to develop a fruit fly optimization algorithm to improve the accuracy and mobile node localization for environmental monitoring.

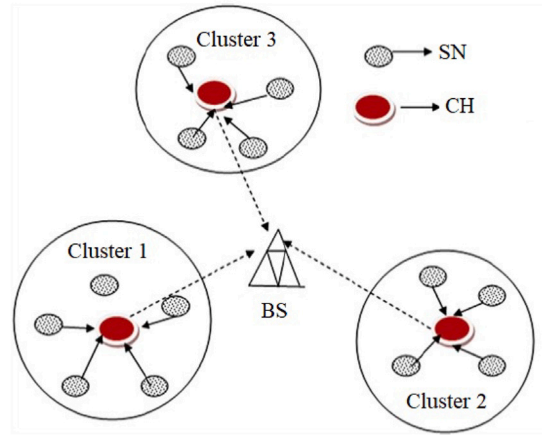


Fig. 5. IoT clustering.

2.2. Clustering and routing in IIoT

Routing is a difficult task in the IIoT due to the intrinsic properties that distinguish it from cellular networks. IIoT employs a large number of SNs, making the use of universal addressing techniques impractical. As a result, traditional IP protocols may not be appropriate for the IIoT. In terms of processing, memory capacity, and energy, SNs are severely hampered. However, IP protocols require resource management and are primarily determined by network architecture and application type. There are two broad categories of IIoT based on network design and protocol operation.

2.3. Cluster-based routing protocol

It is an important subset of sensor networks. This technique has numerous advantages, including scalable architecture, data aggregation/fusion, low energy consumption, low load, and high robustness. The first step in the clustering process is to divide the network into smaller groups known as clusters. Second, each cluster has a CH who is responsible for gathering data from its cluster members, aggregating it, and transmitting it to a BS directly or via other CHs. The BS is linked to the internet to broadcast the incidents. Fig. 5 depicts the architecture of the cluster-based network. Based on network organization, a hierarchical routing technique is developed. In general, it is referred to as cluster-based routing, with the primary focus on energy-efficient communication.

2.3.1. Conventional cluster-based routing protocols

Yadav, Chithaluru, Singh, Albahar, et al. (2022) proposed the adaptive energy clustering protocol, which allows clustering by rotating the CHs during the transmission phase. Clustering became the most prominent traditional clustering protocol used as a benchmark for testing newly developed clustering algorithms by rotating the CHs during the transmission phase (Chithaluru & Prakash, 2018; Chithaluru, Prakash, & Srivastava, 2018; Chithaluru, Tiwari, & Kumar, 2019; Heinzelman, Chandrakasan, & Balakrishnan, 2002; Qin, Li, & Shen, 2018; Zhang, Zhang, Zhang, Dong, & Zhang, 2018). The LEACH procedure is divided into two stages: the setup phase and the steady-state phase. During the first phase, each SN chooses itself as CH with a probability of $\alpha_i(y)$ as given in Eq. (11).

$$\alpha_i(y) = \frac{n}{\eta - n \times (r \times \mod \frac{\eta}{n})}; \quad i \in G_c \qquad (11)$$
$$= 0 \qquad otherwise$$

where c is the number of rounds, n is the expected number of clusters, η is the number of SNs, and G_c is the number of SNs that failed to become CH in the previous $c \times \frac{\eta}{n}$ iterations.

The disadvantages of LEACH are listed below:

- The remaining energy of SNs is not taken into account when selecting CH.
- The probability criterion for becoming a CH is based on the assumption that each SN has a similar primary energy. However, this is not the case in heterogeneous networks.
- Because of the probability rule, the number of CH nodes in this protocol is constant.

Prakash, Chithaluru, Sharma, and Srikanth (2019) proposed modifying the LEACH's threshold equation $\alpha_i(y)$ in conjunction with the SNs' term energy. The problem with LEACH-E is that it ignores the SN's location in CH selection and ignores non-uniformity in CH distribution. Similarly, another proposal includes two factors in the threshold function $\alpha_i(y)$ as emphasized. The first focuses on selecting the CH among a dense distribution of SNs. The second is concerned with the distance between SNs and CHs.

Prakash and Chithaluru (2021) proposed the Hybrid-Energy Efficient Distributed Clustering-Approach (HEED), in which CHs are chosen based on residual energy, and nodes are assigned a probability of becoming a CH based on Eq. (12).

$$CH_{Prob} = \varsigma_{Prob} \times \frac{\epsilon_{res}}{\epsilon_{max}} \tag{12}$$

where, ς_{Prob} is the initial % of CHs, ϵ_{res} is SN's remaining energy and ϵ_{max} is maximum energy of SN's.

If CH_{Prob} is greater than the random number of HEED, the SN is transformed into a CH. Communication costs are reduced by reducing communication overhead. In this case, the network lifetime is extended, and cluster formation is effective. The HEED expends more energy than an issue on the native transmission and CH to BS transmission.

Chithaluru, Singh, and Sharma (2020) developed the Enhanced Distributed-Weight based Energy-efficient Hierarchical Clustering (EDWEHC) protocol. The primary goal of EDWEHC is to improve HEED performance. The improvement is accomplished by creating an optimized intra-cluster topology and balanced cluster sizes. Using Eq. (13), each SN in EDWEHC determines a weight function to become a CH.

$$\omega_w(X) = \frac{\epsilon_{res}(x)}{\epsilon_i(x)} \times \sum_{z} \frac{C-D}{6C} \tag{13}$$

where $\epsilon_i(x)$ is the primary energy of SNs, ε-cluster span to an SN's spatial distance and CH, d is the distance between SNs and adjacent SN z.

Chithaluru, Tanwar, and Kumar (2020) proposed the Enhanced Two-Level (ETL-LEACH) protocol, an improved version of LEACH. It employs self-configuring, randomized cluster formation. As in actual LEACH, a CH gathers information from SNs, but it uses relay nodes to transmit information between CHs and the BS. The benefits of TL-LEACH include random rotation of local clusters and localized coordination, which provide network robustness. Furthermore, this two-level cluster hierarchy scheme results in a low mean transmission distance and a small number of SNs required to send data to distant BS. This method effectively reduces energy consumption. Because of its two-hop data transmission, this protocol is not suitable for wide-area networks.

As proposed by Chithaluru and Prakash (2020), the Energy-Efficient Unequal Clustering (EEUC) protocol employs variable cluster size structural and multi-hop routing. The rotation of CHs in this protocol is based on the remaining energy. As a result, the CHs closer to the BS will expend more residual energy and spend more time transmitting data than the CHs further away. To achieve energy balance, a small number of clusters closer to BS are used. Because of the circular distribution of nodes, this protocol cannot be used for real-time applications.

Ramakuri, Chithaluru, and Kumar (2019) used the K-means algorithm to perform IIoT clustering, and Xu et al. (2012) developed an indoor k-means technique. Using Eq. (14), this method employs K-means to determine the optimal number of clusters.

$$\kappa_{opt} = \sqrt{\frac{\eta}{2\pi} \times \frac{\epsilon_{fs}}{\epsilon_{mp}}} \times \frac{\chi}{D^2} \tag{14}$$

where χ is the network diameter, and ϵ_{fs} and ϵ_{mp} are the RF amplifier energy dissipation model's energy constants. Younis and Fahmy (2004) proposed a centralized and distributed k-means algorithm, with the former running on the BS and the latter on SNs. According to the simulation results, the processing time of the distributed technique is significantly shorter, but there is no difference in energy consumption.

The PSO algorithm aids in the organization of the network into the same number of clusters (Ding, Holliday, & Celik, 2005). This algorithm has two major stages: the first allocates the network into same-sized clusters, and the second uses recursive PSO to select the best CH. When compared to the k-mean algorithm, this method made a successful attempt to form an equal cluster; however, the communication distance for k-means is less due to the center position of CH. Furthermore, because PSO is recursive, the computational cost is higher. Loscri, Morabito, and Marano (2005) used PSO to find the best cluster route by calculating distance without considering the remaining energy of SNs. Eq. (15) computes the objective fitness function Δ.

$$\Delta = \sum_{j-1}^{k} \sum_{i=1}^{ij} (D_{ij}^2 + \frac{D_j^2}{n_j}) \tag{15}$$

where D_{ij} is the distance between SN_i and its CH_j, D_j is the distance between CH_j and BS, and n_j is the number of SNs in cluster j. The PSO algorithm has been thoroughly tested by varying the inertia weight and the acceleration constant (Chithaluru, Kumar, Singh, Benslimane, & Jang, 2021). Behera, Mohapatra, Samal, Khan, Daneshmand, and Gandomi (2019) used a centralized PSO algorithm to reduce intra-cluster spatial distance and increase energy utilization. Behera, Mohapatra, Samal, and Khan (2019) proposed a PSO enhancement to address the uneven clustering issue. They devised the fitness based on gearbox distance. However, the overall node's remaining energy and lifetime are not taken into account. Sasikumar and Khara (2012) used a distributed PSO-based algorithm to choose two CH for each cluster, a Master CH (MCH) and a Slave CH (SCH). The MCH must collect, aggregate, and send the data to the CH, who must then pass the data to the BS. The PSO selects the best MCHs and SCHs based on an objective function that includes cluster communication distance and residual energy computed according to Eq. (16).

$$\delta = \in \times \frac{CH_{egy}}{\varepsilon_{egy}} + (1-\varrho) \tag{16}$$

where ε_{egy} is the cluster's total energy, ϱ is a constant, and CH_{egy} is the cluster's CH energy.

Tillett, Rao, and Sahin (2002), proposed PSO and Cuckoo-embedded methods to reduce energy consumption and transmission distance in the IIoT. The objective function is computed using Eq. (17), Eq. (18), and Eq. (19).

$$\sigma = c \times \sigma_1 + (1-c) \times \sigma_2 \tag{17}$$

$$\sigma_1 = \sum \frac{\epsilon(n)}{\epsilon(CH)} \tag{18}$$

$$\sigma_2 = \sum D(n, CH) \tag{19}$$

where σ_1 denotes the total SNs energy to total CH energy ratio. The distances from all SN to the CH are represented by σ_2. The proposed algorithm is designed to accelerate convergence. In comparison to the Genetic Algorithm (GA) and the PSO, the results demonstrated faster convergence. In comparison to LEACH, the lifetime is significantly longer. Nonetheless, the distance between the CHs and the relay node is not taken into account in this method; however, the global distance between the SNs and their CHs is taken into account. This is more complicated because it runs PSO multiple times to achieve the best result. Guru, Halgamuge, and Fernando (2005) proposed a protocol using PSO with the fitness function estimated as Eqs. (20)–(23).

$$\sigma = \emptyset_1\sigma_1 + \emptyset_2\sigma_2 + (1 - \emptyset_1 - \emptyset_2)\sigma_3 \tag{20}$$

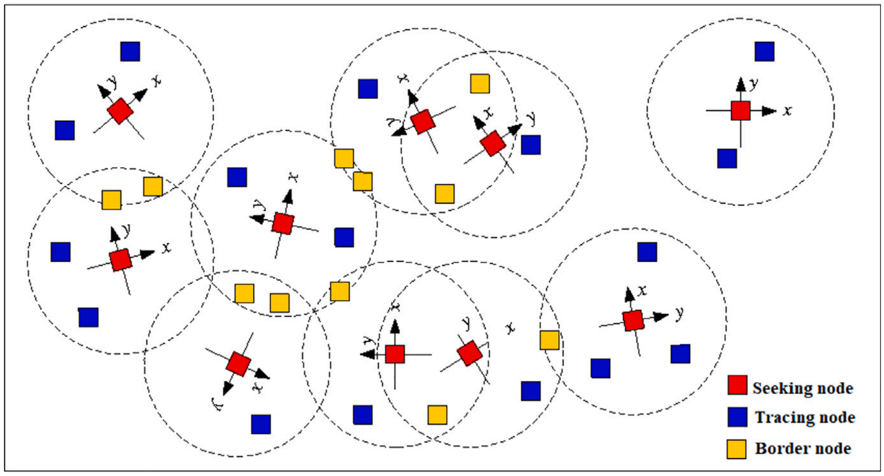


Fig. 6. System model of ICSO protocol.

$$\sigma_1 = \sum \frac{d(n,a)}{\eta_{r-k}} \tag{21}$$

$$\sigma_2 = \sum \frac{\epsilon(a)}{\epsilon(cluster)} \tag{22}$$

$$\sigma_3 = \frac{1}{\psi} \tag{23}$$

$d(n,a)$ represents the distance between the η and the ψ, and η_{r-k} represents the number of SNs belonging to the ψ. $\epsilon(a)$ and $\epsilon(cluster)$ are the energy of the chosen CH and the total energy of the cluster, respectively. The headcount associated with the ψ is represented by ψ. The result is more effective for PSO.

3. Proposed method

This section describes a method of localization that makes use of RSSI data, deployment data, and distance. Following deployment, the SN locations are determined by RSSI value, and the SNs transmit the information to their neighboring nodes to exchange radio signal strengths. In this paper, a new heuristic algorithm for minimizing localization error is proposed. The proposed method builds the probability function of unknown nodes to a correct position using both deployment information and the adjacent node's distance approximation by RSSI. The node placement pattern is illustrated in Fig. 6.

3.1. Range based method

The issue in the ranging phase is calculating the distances between unknown SNs and beacon nodes to improve location accuracy. Many researchers provided direction to eliminate channel interference and reduce RSSI-ranging errors. According to the free-space model, the power received at the receiver (α_ρ) can be expressed as shown in Eq. (24).

$$\alpha_\rho = \alpha_\tau (\frac{\Lambda_0}{4\pi D})^n \tag{24}$$

where α_τ represents the transmitter's signal power; D represents the distance between the transmitter and receiver (in meters); and n is the path loss constant. In practice, due to the number of paths, and noise, the calculated received power differs from the free-space propagation model. Various studies have revealed the common logarithm of the distance path model used in Eq. (25)

$$\alpha L(D) = \alpha L_0 + 10 n log_{10}(\frac{D}{D_0}) + \Theta \tag{25}$$

where D_0 denotes the reference distance; Path loss at the reference distance is represented by αL_0. Gaussian random variable (Θ). The RSSI of the receiver node is calculated using Eq. (26);

$$RSSI = \alpha_\tau - \alpha L(D) \tag{26}$$

where, α_τ denotes signal power of transmitter; reference signal strength at distance D_0 is 1.5 m is θ, then

$$RSSI = \theta - 10 n log_{10}(D) + \Theta \tag{27}$$

Eq. (27) is used to calculate the distance D in terms of RSSI.

3.2. ICSO in localization

The ICSO algorithm is a novel heuristic algorithm that is inspired by cat behavior. It has two sub-models, "seeking mode" and "tracing mode", that mimic cat behavior.

- **Seeking mode:** Ensure a copy of the cat and record its current location. The fitness value for each target point is calculated using Eq. (28). Select the best candidate locations for replacing them with accurate measurements.

$$\sigma_{val} = \frac{(|\sigma_i - \sigma_b|)}{\sigma_{max} - \sigma_{min}} \tag{28}$$

  where σ_i represents current fitness and σ_b represents optimal fitness. If the goal is to reduce the size of the solution, σ_b is equal to σ_{min}; otherwise, σ_b is equal to σ_{max}.
- **Tracing mode:** These sub-model cats represent the praying cat's behavior. To keep the velocities for each cat direction up to date. Check for in-range velocity rather than the maximum limit. To update the position of each cat, employ Eqs. (29) and (30).

$$\varepsilon_{id} = \varepsilon_{id} + c \times s(S_{best} - s_{id}) \tag{29}$$

$$s_{id} = s_{id} + \varepsilon_{id} \tag{30}$$

  where s_{id} represents the cat's actual position, s_{best} represents the cat's local best position, ε_{id} represents the cat's velocity in a M-dimensional state space, c represents a random value (0–0.5), and velocity constant (1–2).

The ICSO technique is used in the design of Finite Impulse Response (FIR) filters, as highlighted in Kulkarni and Venayagamoorthy (2010). The ICSO algorithm, like other adaptive algorithms known as cats, is designed to generate random approaches.

Algorithm 1 ICSO Algorithm

Start procedure
Objective function $\sigma(s)$; where $s = (s_1, s_2, ..., s_n)$
Create population of n cats and Self Position Consideration (SPC), s_i; where $i = 1, 2...n$;
while $\vartheta_{max} > \vartheta$ **do**
  for Estimate the fitness value of each cat **do**
    Best fitness value has been updated $a_{best} = s_{best}$;
    Maintain a new current value s_{best};
  end for
Choose the cat with the highest fitness value;
  for $i = 1{:}n$ **do**
    if SPC $\neq 0$ **then**
      Apply seeking mode;
    else
      Apply tracing mode;
Estimate cat velocity as per Eq. (29);
Position of the cat should be updated as per Eq. (30);
    end if
  end for
end while
Analyze the final results;
End procedure

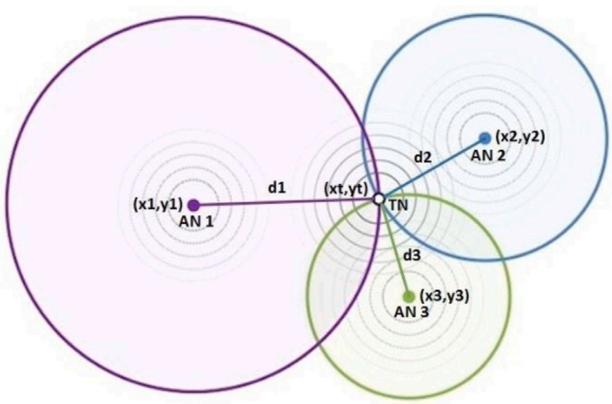


Fig. 7. Trilateral positioning method.

3.2.1. Iterative node localization

The distributed technique in this proposed approach is based on the range, which is used to estimate the position coordinates of unknown nodes by minimizing the objective function. For localization, the following steps are employed:

- In a simulated environment, randomly place the M sensors nodes (target node) and η_q beacons in a sensor field with a communication distance C.
- The beacon nodes are aware of their coordinates.
- Using Eq. (31), estimate the actual distance D_i between the SNs and anchors.

$$D_i = \sqrt{(s_i - s_{i'})^2 + (t_i - t_{i'})^2} \tag{31}$$

- Apply ranging methods to allot the measured distance by the anchors. This is accomplished by incorporating noise into the actual distance $\hat{D}_i = D_i + n_i$, where t_i is a uniformly distributed random value in the frequency band.
- To localize an unknown target node, a minimum of three beacon nodes should be required within the communication radius of the target (known) node. The trilateral method identifies the coordinates of the three beacon nodes P, Q, R, and S, as well as the distances to the target node D_i. The target node coordinates are calculated using trigonometric laws. The distance measurements of three or more beacon nodes are used in this approximation scheme to reduce the error between the estimated and actual distances. The calculation process is depicted in Fig. 7.

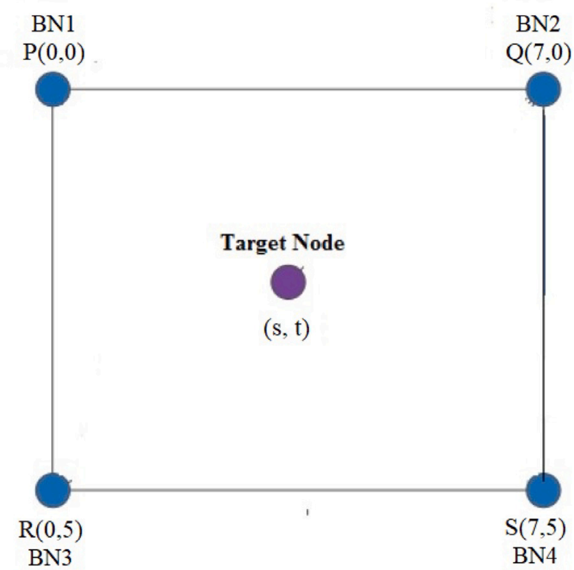


Fig. 8. ICSO experimental localization.

- To determine the target node position, each localized node is initialized with the geometric center of the beacon nodes that are within the transmission area, as per Eq. (32).

$$(s_c, t_c) = (\frac{1}{\eta}\sum_{i=1}^{\eta} s_i, \sum_{i=1}^{\eta} t_i) \tag{32}$$

where η_q denotes the total number of beacon nodes within the target node's communication distance.

- The bio-inspired methods aid in determining the coordinates (s, t) of the unknown SN, reducing error. The objective function, which represents the error function, can be mathematically expressed in Eq. (33).

$$\sigma(s, t) = \frac{1}{\eta_p}\sum_{i=1}^{\eta_p}(D_i - \hat{D}_i)^2 \tag{33}$$

where D_i is the actual distance between unknown and anchor nodes and (s_i, t_i) is the relative anchor positions; (s_m, t_m) represents the position occupied by SNs; and η_p is the number of anchors within SN communication accessibility.

- The algorithms return the closest values of the coordinates (s_m, t_m) with the least amount of error. Similarly, the algorithm finds the next unknown node within the coverage range.
- The already-located SNs are removed from the unknown target list and serve as anchors in the next round of operations.
- Based on its estimated coordinates, the number of localize nodes η_l is used to calculate the localization error. According to Eq. (34), (s_i, t_i) is the mean of squares of distances between actual and estimated node locations. ICSO and PSO estimate (s_i', t_i'), $i = 1, 2, \ldots, \eta_l$.

$$\epsilon_l = \frac{1}{\eta_l}\sum_{i=1}^{\eta_p}((s_i - s_i')^2 + (t_i - t_i')^2) \tag{34}$$

until all unknown (target) nodes are located; otherwise, no new nodes will be added to the network. The values of $\eta_{\eta l}$ and ϵ_l, where $\eta_{\eta l} = \eta_a$ - η_l is the number of nodes that could not be

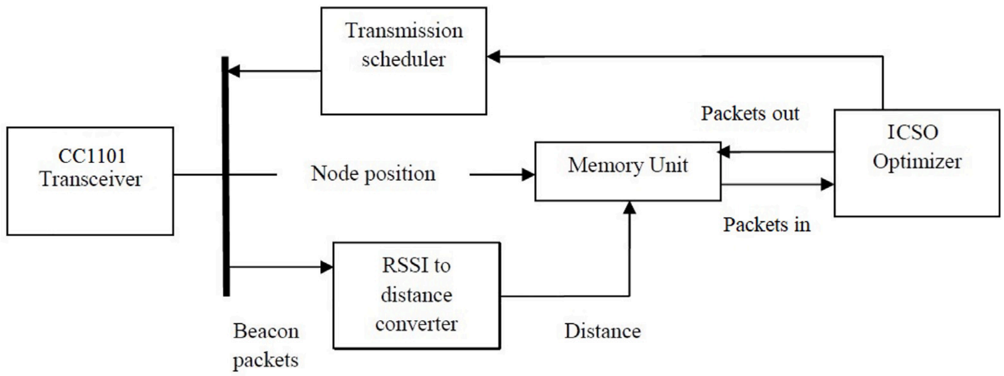


Fig. 9. Proposed localization model.

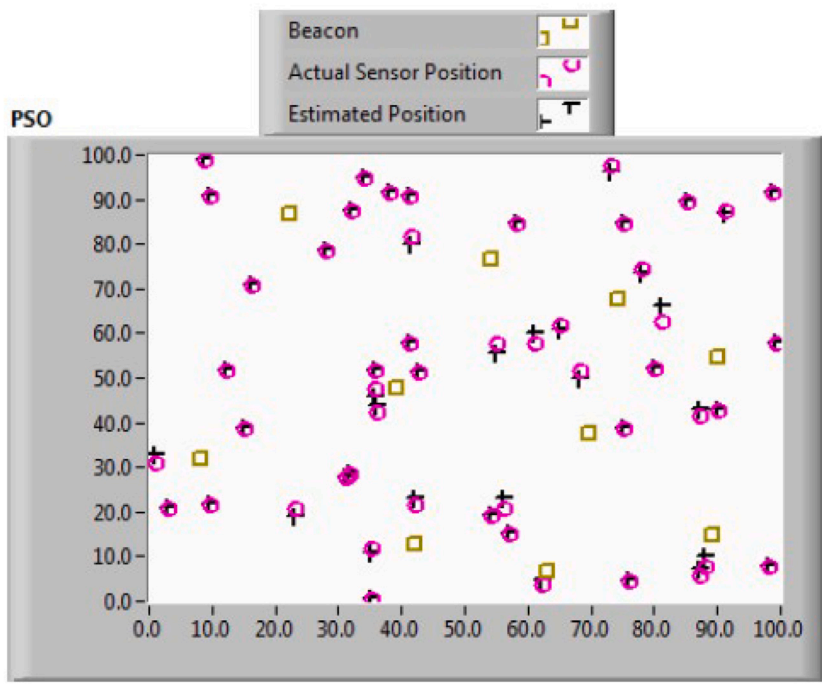


Fig. 10. PSO estimated locations (η = 20).

localized and ϵ_l is the number of nodes that could not be localized, clearly show the performance of the localization algorithm. Lower η_{nl} and ϵ_l values indicate better results.

3.3. Real-time implementation for indoor node localization

3.3.1. Preliminaries

An IIoT node positioning method based on Wi-Fi technology is proposed in this scenario. The node of the localization network is the Texas Instruments CC1101 low-power transceiver. The RSSI values of the target nodes within the communication region can be detected by the CC1101 transceiver with PIC16F887-I/P. Fig. 8 depicts the experimental setup. The RSSI values received from its beacon nodes are used to locate the target node in the center. The gateway node serves as the foundation for IIoT. To begin, all beacon nodes receive the positioning information signal sent by the gateway node. The beacon nodes then send signals to the target node, which then sends data to the gateway node. The RSSI values used for position estimation are extracted by the gateway node using MATLAB. The ICSO positioning algorithm is run by MATLAB, and the value of the SNs position of (unknown) target SNs is provided as coordinates.

The experiment was carried out in a 7 m × 5 m indoor area. As shown in Fig. 8, 4 beacon nodes were used in this experiment: P, Q, R, and S. These beacon nodes' coordinates are (0, 0), (7, 0), (0, 5), and (7, 5), respectively. The target node, whose coordinate (s, t) is unknown, cannot move beyond the area PQRS's dimension. Three beacon nodes move the target node to different possible locations to estimate their position awareness.

3.3.2. Localization system model

Fig. 9 depicts the model of the localization system. It includes modules such as a transceiver, RSSI to distance converter, ICSO optimizer, storage unit, and communication scheduler. The beacon nodes are placed in known locations and communicate their location to the target node. The communication scheduler then requests that the beacon nodes transmit beacon data packets containing their estimated positions. Furthermore, the RSSI value from beacons is measured using the RSSI-distance converter and then converted into distance. The ICSO optimizer then uses Eq. (33) to determine the SN's predictable position. A gateway node receives packets from deployed SNs and sends them to the BS for calculation.

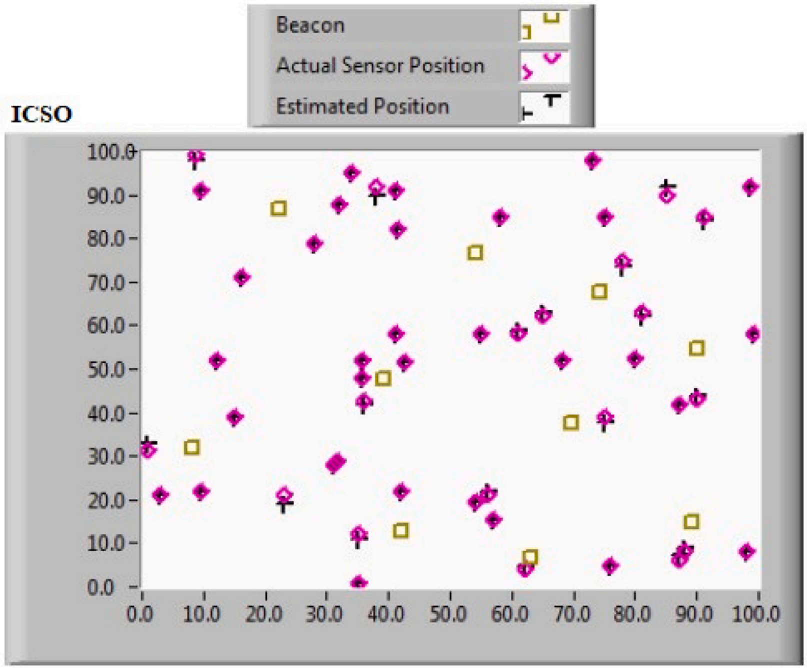


Fig. 11. ICSO estimated locations (η = 20).

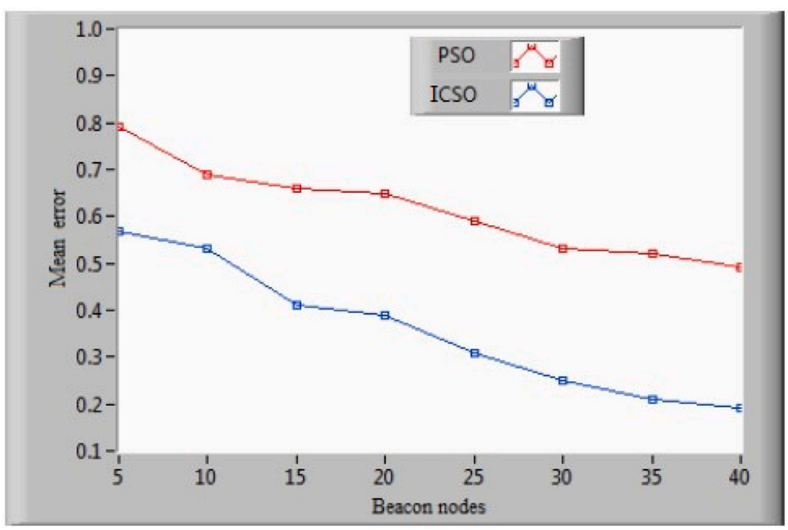


Fig. 12. Beacon node density Vs. error.

- **Transceiver system:** Each SN is equipped with a transceiver and a microcontroller unit. The transceiver unit will receive beacon messages from neighboring SNs and send them to the next stage for distance conversion. In this case, the hardware included a low-power CC1100, a 1.2 GHz transceiver, and a microcontroller PIC16F887-I/P for data processing.
- **Transmission/Reception schedule:** To avoid data collisions during transmission and reception, a transmission scheduler has been developed to implement the Media Access Control (MAC) protocol. For ease of use, a simple technique was used. When an SN becomes active, it begins polling nearby SNs for beacon messages.
- **RSSI to distance converter:** The CC-1101 transceiver includes an RSS indicator, which can be output via an analog output pin. Eq. (35) is used to calculate the RSSI to distance.

$$\hat{D}_i = 10^{\frac{D_0 - RSSI}{2n}} \tag{35}$$

where $\hat{D}_i$ denotes the distance, D_0 denotes the received RSSI value for a distance of 1 m, and n denotes the path loss exponent. Before the system can be deployed, D_0 and n are obtained. The power levels are converted into the distance by MATLAB.

- **ICSO optimizer:** Using Eq. (33), the ICSO optimizer is used to determine the SN's predictable position by combining deployment and inter-node distance data. For optimization, the ICSO algorithm is used to find the SN's position. The ICSO-optimizer uses a random weight to vary the velocity at which the cat moves ahead of its A_{best} and W_{best}. In this instance, each cat has two members that represent the s and t coordinates of SNs.

The temperature sensed from the environment is transmitted to the BS by all four beacons (BN1, BN2, BN3, and BN4). The packets contain the node ID, duration, voltage of the battery, and RSSI value of the node. A serial port is used to transfer data from SNs to BS.

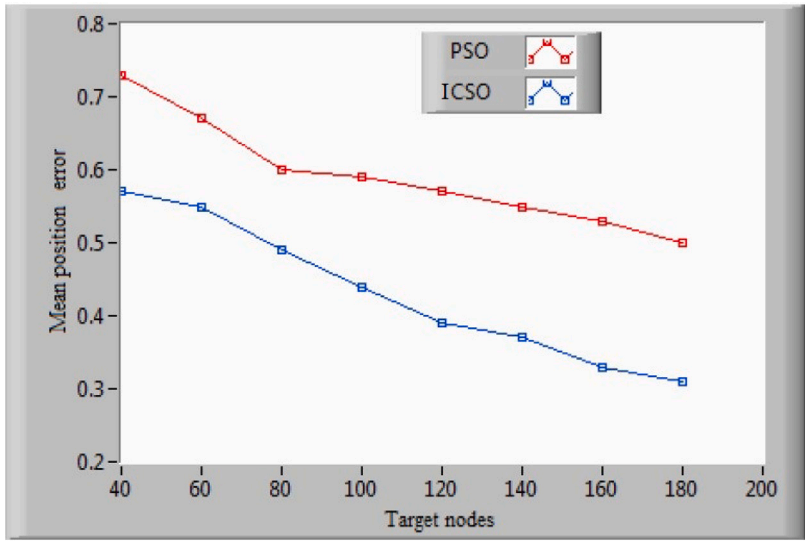


Fig. 13. Target node density Vs. error.

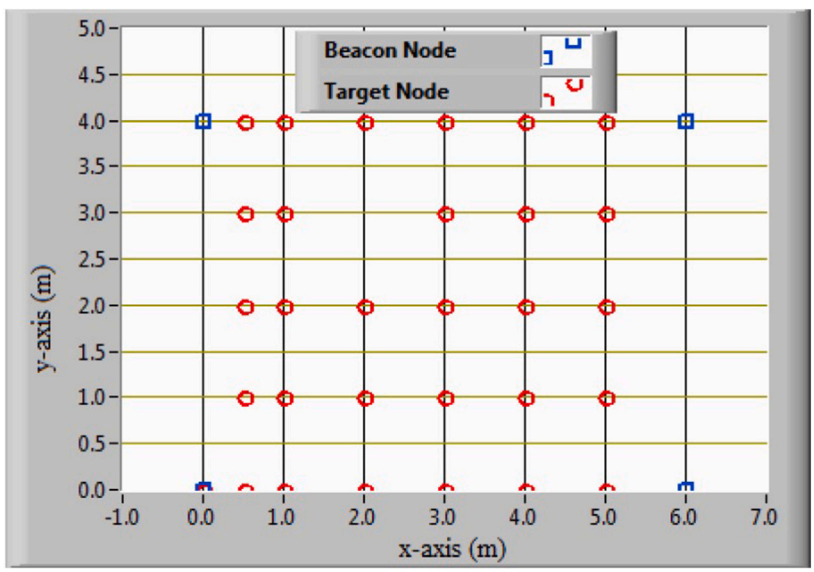


Fig. 14. Sensor field deployment (beacon vs. target node).

4. Results and analysis

4.1. Simulated environment

The IIoT localization in simulation is performed on a Windows computer using MATLAB. For the preliminary investigation, 50 target nodes and 10 anchor nodes are organized in a random two-dimensional field with 1000 × 1000 square units. Every anchor has a 50-unit communication radius. The experiment is carried out for approximately 30 localization trials with Gaussian additive noise of $\alpha_n = 3$ and $\alpha_n = 6$.

Five trial runs out of 50 are used in the detailed study on localization using all three algorithms. Because the results of different trial runs are not identical due to the stochastic nature of the algorithm, the average results of experiments conducted are summarized in Table 3. The number of nodes localized and the localization error are used to analyze performance metrics. ICSO values are lower than PSO values, as evidenced by ICSO's superior performance. Similarly, the Computation Time (τ_c) required for ICSO is significantly less than that required for PSO. The algorithms' position estimation for beacon and target nodes is depicted in Figs. 10 and 11. The proposed new algorithm is also evaluated for various aspects of localization accuracy, such as the effect of additive Gaussian noise, beacon, target node densities, and communication radius. The network parameters for simulation are listed in Table 2.

Effect of Gaussian noise

When the percentage of noise (α_n) in distance measurement is reduced from 3 to 6, it appears to reduce the localization error for all three algorithms. The Gaussian noise effect is tabulated in Table 4.

Effect of beacon node density on localization error

When there are more beacon nodes to localize the unknown nodes, the number of localized nodes increases; additionally, beacon $\eta \geq 4$ is required to locate the position of unknown nodes. As shown in Fig. 12, the percentage of localized nodes is proportional to the number of beacon nodes.

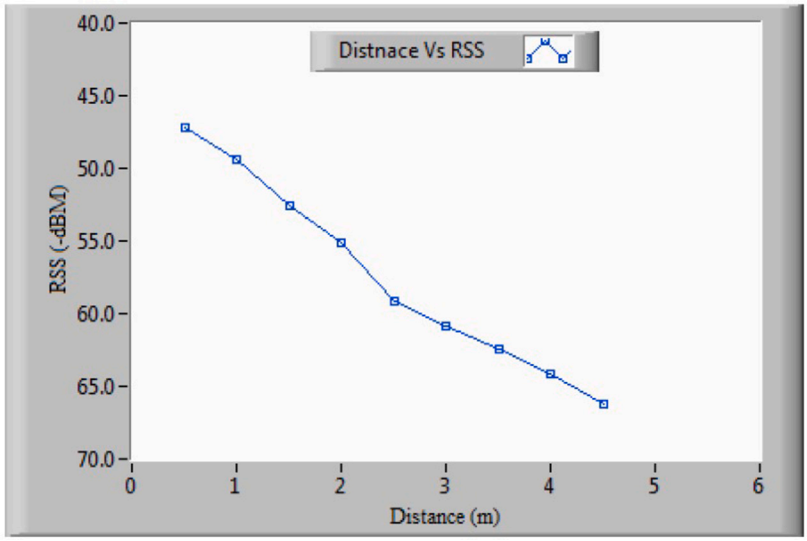


Fig. 15. Beacon RSSI Vs. distance.

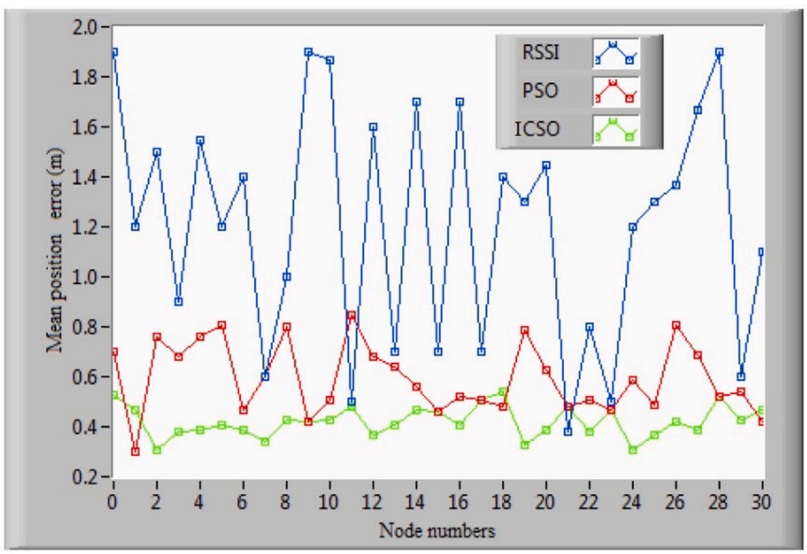


Fig. 16. Position error computed by RSSI Vs. PSO Vs. ICSO.

Table 2
Test-bed parameters of network.

Parameter	Value
Sensor field	1000 × 1000 sq m
BS location	(400–500), (400–500)
Number of SNs	100
Energy of SNs	1.5 J
ϵ_{ele}	50 nJ/bit
ϵ_{mp}	0.0013 pJ/bit/m4
ϵ_{fs}	10 pJ/bit/m2
ϵ_{da}	5 nJ/bit/message
D_{max}	1000 m
D_0	250 m
Packet length	2400 bits
Message size	1000 bits

Table 3
Summary of PSO and ICSO results.

Trial	Iteration	PSO			ICSO		
		$\eta_{\eta l}$	ϵ_l	$\tau_i(x)$	$\eta_{\eta l}$	ϵ_l	$\tau_i(s)$
Trial = 1	20	49	0.380	1.65	50	0.315	1.60
η = 50	40	49	0.416	1.69	44	0.423	1.16
M = 20	60	48	0.648	1.80	47	0.319	1.45
ρ = 50	80	50	0.723	1.84	49	0.305	1.47
	100	44	0.489	1.73	50	0.278	1.13
Trial = 2	20	97	0.724	3.56	100	0.383	3.26
η = 100	40	98	0.767	3.41	100	0.423	3.15
M = 30	60	98	0.679	3.67	100	0.435	3.25
ρ = 50	80	100	0.687	4.12	100	0.515	3.19
	100	98	0.691	4.671	100	0.525	2.27

Effect of number of nodes

The number of nodes is increased from 50 to 200 to better understand the performance of localization. Fig. 13 depicts the effects of target node density as the density of localized nodes increases.

4.2. Experimental performance

Compare the performance of the ICSO algorithm and the RSSI-based localization algorithm in terms of the localization error. The test was carried out at 30 different target node positions within the PQRS area.