International Journal of Computer Networks and Applications (IJCNA)
DOI: 10.22247/ijcna/2024/41 Volume 11, Issue 5, September – October (2024)

RESEARCH ARTICLE

Integrating MANETs and Hybrid Deep Learning for Enhanced River Water Quality Monitoring

D. Manopriya
Department of Computer Science, Sri Krishna Adithya College of Arts and Science, Sugunapuram, Coimbatore, India.
dmanopriya56@gmail.com

M. Manimaran
School of Computing Science Engineering and Artificial Intelligence, VIT Bhopal University, Bhopal, Madhya Pradesh, India.
mani.professor@gmail.com

Received: 15 July 2024 / Revised: 04 October 2024 / Accepted: 11 October 2024 / Published: 30 October 2024

Abstract – Rivers play a critical role in supporting ecosystems and human activities. However, water quality in rivers is increasingly threatened by various pollutants. The growing threats to river ecosystems demand immediate and effective solutions to monitor and mitigate pollution. Traditional water quality monitoring methods have several concerns such as high cost and difficulty to deploy in remote areas. The necessary actions are needed to control river pollution. Mobile Ad-hoc Networks (MANETs) offer an effective solution for water quality monitoring in real-time based on their infrastructure-less, self-configuring, and scalable nature. In this work, the integrated solution is proposed for water quality monitoring. It includes new clustering techniques and deep models for effective communication prediction. For reliable communication, a new clustering protocol is proposed by considering multiple clustering metrics. The Cluster Head (CH) is selected using the metaheuristic optimization algorithm of the Walrus Optimization Algorithm (WOA). To achieve higher accuracy in prediction, a new hybrid deep learning (DL) model combines Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) models. Simulation results show the proposed approach is better in terms of delay, packet delivery ratio (PDR), and energy consumption when compared to other techniques.

Index Terms – MANET, Water Quality, WOA, Clustering Protocol, Hybrid Deep Learning Model, Convolutional Neural Networks.

1. INTRODUCTION

Rivers play major roles in ecosystems and human societies. It is supported by giving water for drinking, agriculture and recreational activities [1][2]. Today, the water quality is extremely vulnerable to many pollutants and their harmful effects. The monitoring of water quality is very essential for several reasons: Pollution Detection, Ecosystem Health, Public Health and Agricultural Impact. Early detection of pollutants is vital to prevent environmental degradation and protect aquatic life. Further, the appropriate measures can be taken to maintain water quality and safeguard these vital resources.

Mobile Ad-hoc Networks (MANETs) an infrastructure-less networks which connect wirelessly [3]. The MANET technology is well suited for dynamic and remote environments. The unique behaviour of MANETs are Self-Configuration, Dynamic Topology, Scalability and Decentralization [4]. The nodes in the network automatically discover and configure themselves. It allows flexible and adaptive communication paths. In addition, MANETs can easily scale by adding new nodes to cover larger areas or by replacing faulty ones. These features make MANETs the best choice for collecting and transmitting water-quality data in real-time.

Despite their advantages, MANETs encounter several challenges that can impede their effectiveness in real-world applications. The limited bandwidth restricts data transmission rates when multiple nodes communicate simultaneously. The energy consumption is critical due to the battery-powered nodes. The dynamic topology environment causes routing challenges which require robust protocols. Lastly, data integrity and reliability are compromised due to transient connections. It causes packet loss or corruption in this scenario.

The implementation of MANETs in river water quality monitoring meets essential needs for environmental sustainability. Real-time data collection is enabled through the dynamic configuration of nodes. It allows continuous monitoring from various sensors deployed along the river. In addition, cost-effectiveness is achieved by reducing the need for fixed infrastructure and reducing expenses related to traditional monitoring networks.

In MANET, clustering is the process of grouping nodes based on some criteria. This clustering concept is used to optimise communication and resource management within the network. Clustering is a technique used in MANETs to organize nodes into manageable groups, each managed by a CH [5]. This hierarchical structure optimizes communication and resource management within the network. It reduces the number of communications between nodes and reduces the energy consumption of individual nodes. So, the clustering process improves the efficiency and effectiveness of data collection and transmission in MANET river quality monitoring systems. In water quality analysis, the DL model has received greater attention based on their accuracy and flexibility [6-11]. It can capture complex temporal dependencies and water quality data patterns and reliable predictions.

This work contributed a novel clustering protocol based on the WOA. It improves communication efficiency and reduces energy consumption in MANET-based water quality monitoring applications. This work aims to develop an integrated to monitor water quality using MANETs and advanced DL techniques where the proposed work ideology is given in the following.

Design a reliable and scalable MANET for water quality monitoring. It uses new clustering techniques to optimize network communication and energy consumption.

Propose a novel clustering protocol that considers multiple metrics for cluster formation and selects the optimal CH using the WOA. Develop a hybrid model for water quality prediction with higher accuracy by capturing both spatial and temporal dependencies.

The paper is organized as follows. Section 2 presents water quality prediction techniques relevant works. Section 3 describes the proposed model for prediction. The experimental results carried out in section 4 and section 5 present the summary of the work.

2. RELATED WORK

Wang Jingmeng, et al [12] explored a fuzzy method for TongHui River water prediction. It applies fuzzy memberships and spatial clustering to analyze the water quality parameters. However, the fuzzy-based model may not capture temporal dependencies in the data which could affect long-term predictions.

In their study, Yunchao Jiang et al [13] introduced an artificial neural network (ANN) integrated water quality assessment model in the Yellow River of China. The features are extracted using self-organizing maps. Then, the ANN model is applied for quality categorization. Likewise, the author Z. -l. Hao et al [14] presented a multilayer feed-forward neural network model for quality prediction. The model weights are altered to learn the complex pattern of water quality data.

Z. -l. Hao, et al [15] derived a mathematical model based on an empirical equation to predict water quality in the Indian River. It considers multiple attributes like pH, Oxygen Demand and temperature for accurate prediction. Similarly, A. K. Shukla et al [16] analysed the quality of the Indian River using index mapping. In addition, the geographic information system is used to consider spatial relationships. In [17], the author introduced an Internet of Things (IoT) enabled solution for monitoring water quality in rivers. The collected data from the sensor is updated on the server every fifteen days. The artificial intelligence-based algorithm is applied for quality analysis. Likewise, R. P. N. Budiarti et al [18] presented a Raspberry Pi processor for water quality. However, the system's accuracy is dependent on the quality and maintenance of the sensors which can degrade over time. Similarly, A. L. Lopez [19], proposed an IoT system combined with DL models for water quality prediction. The overall accuracy achieved by the model is only 91.4%. The hybrid model which combines fuzzy rules and Maximum Connection Degree Principle (MCDP) is used by Shuang Zhu et al [20] for Huai River Basin quality detection. MCDP extracts hidden features from the data set for accurate prediction. Results show that the hybrid model attains 92% accuracy on the test data. However, the model requires more training data.

Chellaswamy C et al [21] constructed a river water quality assessment. It includes a CNN and recurrent neural network to learn the temporal patterns of water quality data. By combining the CNN and recurrent network, the proposed model further improves the accuracy by at least 8.8%. However, the model complexity increases and requires more computational resources.

In [22], the author analyzed the performance of a gated recurrent model for water quality analysis. The gated model effectively handles vanishing gradient and overfitting issues compared to other recurrent models. The proposed model reduces the error rates by 12.89 compared to the existing model. But it requires manual tuning of parameters.

Here Gradient Boosting Model (GBM) is used by A. -A. Nayan et al [23]. The water parameters from the year 2012 to 2018 are collected from the Songhua River Harbin Region and processed using GBM models. Compared to random forest models, the GBM model shows higher accuracy in prediction. In [24] , the author proposed an adaptive fuzzy system in the Hei River water prediction. Initially, the membership rules are framed based on input variables. Based on training data, the membership functions are updated. However, the model’s performance is highly dependent on the input data quality and requires complex tuning of membership functions.

RESEARCH ARTICLE

The three-fold model is proposed by A. P. Kogekar et al [25] in the Ganga river water. It includes CNN, LSTM and Support Vector Regression for extracting features individually. Then, the fusion process is applied for final quality prediction.

L. Guo et al [26] proposed an Extreme learning machine (ELM) based river water quality prediction model. Initially, the data is pre-processed using principal component analysis. Then, the extracted components or features are further processed by ELM for prediction. In addition, the particle swarm optimization base parameter tuning is used in ELM for accurate prediction. The execution time of the model is higher due to the local minima issues of the swarm optimization algorithm.

The hybrid LSTM model-based water quality prediction technique is proposed by D. Zhang, et al [27]. For preprocessing, the Ensemble Empirical Mode Decomposition (EEMD) is applied which decomposes data based on the signal's intrinsic characteristics. The LSTM model combined with EEMD shows a higher accuracy than sole models. Similarly, In [28], Z. Guo, et al proposed a Bi-LSTM model combined with an attention model. It is used to process the data selectively based on data importance. However, the integration of attention mechanisms increases the model's computational requirements.

A. P. Kogekar et al [29] applied a modified Auto-Regressive Integrated Moving Average (ARIMA) model to process the Ganga River's water quality. ARIMA is a statistical method which uses the concept of regression to process the sequential data. The implementation results show that the ARIMA model suffers from the nonlinearity of data and an error rate of about 45.2%.

H. Sun et al [30] T-S fuzzy neural network model for the prediction of air quality in Yangtze River. A Takagi-Sugeno (T-S) fuzzy model is used to handle nonlinear and complex systems more effectively. This model uses fuzzy if-then rules to create a fuzzy inference system and it typically employs a neural network to optimize and tune the fuzzy's parameters.

P. Siagian, et al [31] introduced a smart system for water quality monitoring system. The sensor node with embedded controllers is used to collect the data from the river and processed in a base station for quality prediction.

In [32], the author proposed an echo state for analyzing water quality in Zhengzhou City. It works based on reservoir computing to deeply learn the features of water quality parameters. The echo state model requires fine-tuning of the reservoir size to prevent overfitting. It increases execution time by 8.9% when compared to previous models.

S. Chopade, et al [33] applied a sensor network system integrated with DL models for water quality analysis. The CNN is used to differentiate the quality of water data.

A., Rajaram et al [34] proposed a MANET clustering protocol using Cellular Automata and African Buffalo Optimization algorithm for CH selection. Likewise, Venkatasubramanian, S et al [35] proposed a CH selection protocol using the Vortex Search Algorithm (VSA). The fitness function of cluster selection is solved using VSA. Hamza, F et al [36] proposed a fuzzy combined Emperor Penguin Optimization (EPO) based clustering protocol for MANET. EPO is applied to solve the CH selection problem. The overall summary of the above works is in Table 1.

Table 1 Literature Summary

Authors	Methodology	Advantages	Disadvantages
Yunchao Jiang et al. [13]	ANN with SOM for feature extraction	Handles large datasets; clustering via SOM	Does not capture temporal dependencies. Average error rate about 24.52%
Z.-l. Hao et al. [14]	Multilayer FNN	Models non-linear relationships	Does not handle sequential data
Z.-l. Hao et al. [15]	Empirical model with multiple parameters	Simple and interpretable. Overall accuracy is 93.5%.	Limited generalizability
A. K. Shukla et al. [16]	Index mapping with GIS	Considers spatial relationships	Computationally expensive
R. P. N. Budiarti et al. [18]	IoT-enabled system with Raspberry Pi	real-time data collection	Sensor accuracy depends on maintenance

A. L. Lopez [19]	IoT + deep learning	Improved feature extraction	High computational cost
Shuang Zhu et al. [20]	Hybrid model with Fuzzy Logic and MCDP	Higher accuracy of 92%	Computationally intensive
Chellaswamy C et al. [21]	Hybrid CNN and RNN	Improved accuracy of 91.8%	High model complexity
A.-A. Nayan et al. [23]	Gradient Boosting Model (GBM)	High accuracy compared to random forest	Sensitive to hyperparameters, overfitting risk
D. Zhang et al. [27]	Hybrid LSTM with EEMD	Improved accuracy with EEMD	Complex and computationally expensive
Z. Guo et al. [28]	Bi-LSTM with an attention mechanism	Selectively focuses on relevant data	High computational cost
P. Siagian et al. [31]	Sensor network with embedded controllers	Real-time monitoring	Requires robust infrastructure
S. Chopade et al. [33]	Sensor network + DNN	Captures complex patterns	Requires large labeled datasets, prone to overfitting

2.1. Research Gap

From the literature, significant advancements have been made in water quality monitoring through various DL models. However, several gaps remain in the existing literature. Many studies concentrate on specific algorithms but fail to integrate MANETs for real-time monitoring which limits their effectiveness in dynamic and remote environments. Additionally, current clustering protocols neglect the necessity for multi-metric optimization in cluster head selection. It reduces overall clustering performance.

3. PROPOSED CLUSTERING AND QUALITY PREDICTION MODEL

The proposed method consists of both static and mobile units including UAVs and autonomous boats to predict water quality. Static nodes are strategically placed along riverbanks to continuously measure various parameters. The mobile nodes cover areas outside the reach of static sensors.

The network is organized into clusters with optimal CHs selected using the WOA to ensure efficient data aggregation. Data is transmitted to a base station for preliminary processing before being sent to a central server where advanced analysis is conducted using hybrid CNN and LSTM to learn from both historical and real-time data that attain accurate water quality predictions. Figure 1 shows the overall proposed architecture.

3.1. System Architecture

This architecture consists of different key elements. It includes sensor nodes, cluster heads, a base station, and a central server for data processing and analysis.

3.1.1. Sensor Nodes

Sensor nodes are crucial for collecting water quality data. It has static and mobile sensor nodes. Static sensor nodes are deployed at fixed locations along the riverbank and at key points within the river. It continuously monitors pH, temperature, dissolved oxygen and turbidity levels in water. On the other hand, mobile sensor nodes used for an autonomous boat equipped with sensors move along the river to collect data from areas not covered by static nodes.

3.1.2. Clustering and CHs

Initially, the clusters are formed using K-Means Clustering. It partitions the nodes into K clusters by assigning each node to the nearest centroid based on the distance and iteratively updating the centroids until convergence. After clustering, the optimal CH is selected using WOA. The workflow of the proposed approach is given in Figure 2.

3.1.3. Base Station

The base station is strategically located on the riverbank. It serves as the main data collection hub. The main function of a base station is receiving aggregated data from CH and mobile nodes. The base station handles initial data processing and forwards the processed data to the central server for detailed analysis.

International Journal of Computer Networks and Applications (IJCNA)
DOI: 10.22247/ijcna/2024/41 Volume 11, Issue 5, September – October (2024)

RESEARCH ARTICLE

3.1.4. Central Server

It is used to process and analyse advanced data. It receives pre-processed data from the base station and performs further cleaning and preprocessing to handle missing values, noise, and normalization. In this work, a hybrid DL model is used to analyse the data and predict water quality parameters. The hybrid model consists of CNN and LSTM layers to process the sequential water quality data. The model is trained using historical and real-time data to learn patterns and make accurate predictions.

3.1.5. Data Flow and Communication

Initially, sensor nodes (both static and mobile) continuously collect water quality data. The mobile nodes move along the river to cover areas not monitored by static nodes and relay data from static nodes in their vicinity. The collected data is then aggregated at the cluster level, with CHs collecting and aggregating data from their member nodes. The aggregated data is transmitted to the base station using multi-hop communication if necessary. The base station preprocesses the data and forwards it to the central server, where DL models analyze the data for quality prediction.

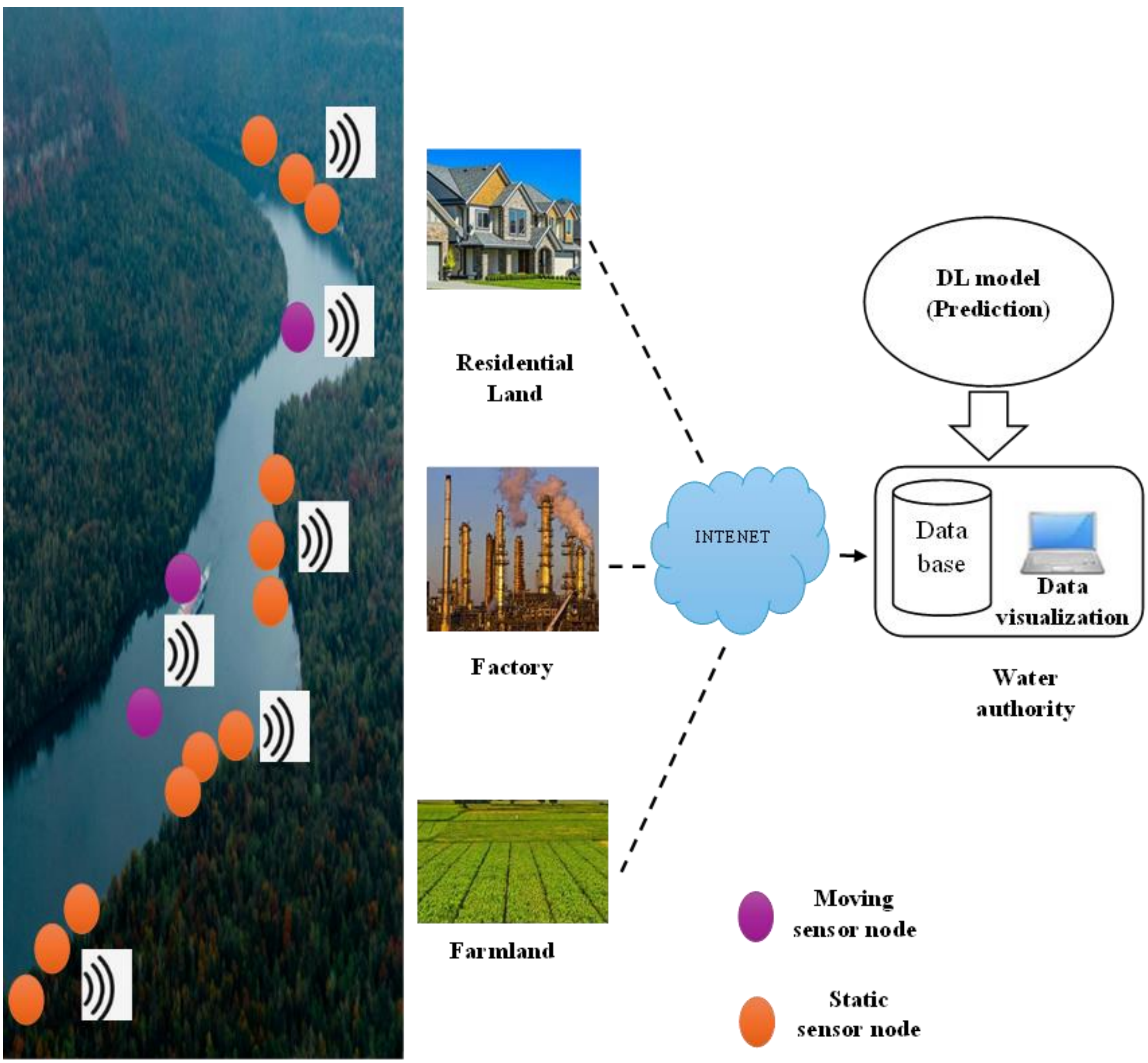

Figure 1 Proposed Architecture

RESEARCH ARTICLE

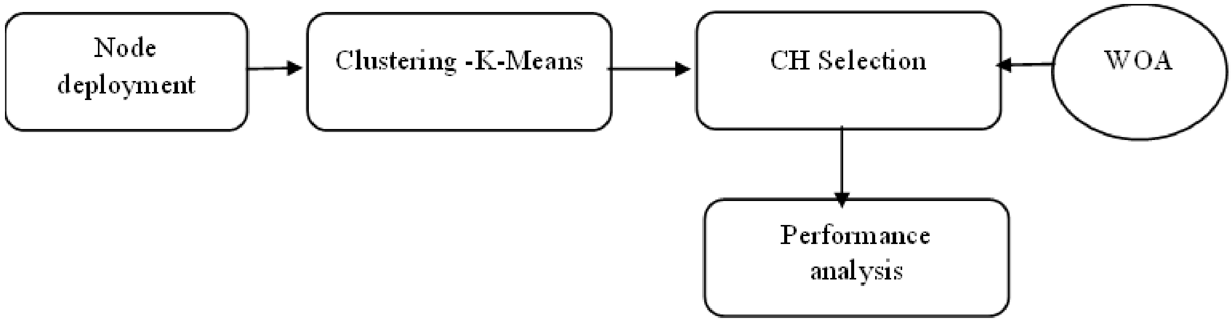

Figure 2 Proposed Clustering and CH Selection

3.2. Walrus Optimization Algorithm (WOA)

The metaheuristic optimization algorithms are used for solving complex problems [37]. WOA is inspired by the behaviour of Walrus in the Arctic Ocean for their survival [38]. It has unique behaviours of roaming, escaping, and fighting predators. This behaviour is motivated to frame the optimization model for solving real-time problems. Unlike traditional optimization techniques, WOA's unique approach minimizes the risk and enables it to explore a broader solution space. It includes three stages: feeding strategy, roaming and escaping

3.2.1. Feeding strategy

Walruses have a diverse diet but prefer benthic bivalve molluscs like clams. Their feeding involves searching the sea floor with energetic motions and sensitive vibrissae (whiskers). The strongest walrus, identified by its tusks, leads the group to find food, akin to guiding the search for the best solutions. It can be mathematically modelled as follows (Eq. 1 & Eq. 2):

$$P_{i,j}^{P1} = P_{i,j} + rand\left(M_{i,j} - I_{i,j}p_{i,j}\right) \quad i=1,2..N \tag{1}$$

$$P_i = \begin{cases} P_{i,j}^{P1}, \ P_i^{P1} < F_i \\ P_i \ , ELSE \end{cases} \tag{2}$$

Where, P_i^{P1} indicates ith walrus newly generated location, rand denotes random number in [0, 1]. M indicate the one member, j is the dimension, $I_{i,j}$are random numbers from the set {1, 2}.

3.2.2. Roaming

This stage is used to explore and exploit for optimization. The walrus moved to the next best location based on its food resources. This migration is mathematically expressed as follows (Eq. 3) :

$$x_{i,j}^{P2} = \begin{cases} x_{i,j} + rand\left(x_{k,j} - Ix_{i,j}\right), (F_k < F_i \\ x_{i,j} + rand\left(x_{i,j} - x_{k,j}\right), ELSE \end{cases} \tag{3}$$

Where, $x_{i,j}^{P2s1}$ is the new location of Walrus.

3.2.3. Escaping

This stage belongs to the exploitation process of optimization. The walrus strategically escaped and attacked the predators and moved to safer locations. This escaping behaviour is mathematically modelled as follows (Eq. 4):

$$x_{i,j}^{P3} = x_{i,j} + (ul_{local,j}^{t} - rand.ll_{local,j}^{t})) \tag{4}$$

Where, $x_{i,j}^{P3}$ is the new location of the walrus, ul and ll is the upper and lower limit of search space.

3.3. CH Selection Process

The CH selection process is considered an optimization problem in this work. For CH selection, the multiple CH selection parameters like node speed, distance and energy. The optimal CH is selected using WOA.

3.3.1. Node Speed

The node speed is the essential parameter in MANET clustering. The node with the lowest speed has minimum distance and has more stability. The node with minimum speed would stay longer time within a cluster. The speed of the node is calculated as (Eq. 5):

$$b_n = \frac{b - b_{min}}{b_{max} - b_{min}} \tag{5}$$

Where b_{min} and b_{max} represents the maximum and minimum speeds of the vehicle. b indicates an average speed.

3.3.2. The distance between the member node and CH

Consider Kth clusters with N_i nodes where i = 1,2,…,K and node's average distance is given in (Eq. 6):

$$\text{dist}_1 = \max_{i=1,2,\dots,K}\left\{\frac{\sum_{j=1}^{N_i}(CM_{ij}, CH_{ij})}{N_i}\right\} \tag{6}$$

An overall energy consumption is computed as follows (Eq. 7):

$$E_{sum} = \sum_{i=1}^{K}(E_{CH}^{i} + \sum_{j=1}^{N_i} E_{mem}^{ij}) \tag{7}$$

The above values of the nodes are normalized into maximum average values. The fitness function for CH selection as a

RESEARCH ARTICLE

function of energy, distance, energy consumption and node speed can be framed as follows (Eq. 8):

$$F = \sigma_1 dist_1 + \sigma_2 dist_2 + \sigma_3 E_{total} + \sigma_4 T_{total} \quad (8)$$

Where, σ_1,σ_2,$\sigma_3$and σ_4 are the positive factor values used to vary the importance of parameters. The summation of all factors is equal to one. Algorithm 1 outlines the proposed CH selection approach.

Step 1: Initialization

Initialize the population of walruses (candidate CH nodes) P_i

Evaluate the fitness F_{i} for each candidate solution P_i

Identify the strongest walrus (best candidate) based on fitness

Step 2: Main loop of WOA

For each iteration do:

For each walrus i in the population do:

Feeding strategy

For each dimension j do:

$P_{\mathrm{i,j}}^{P1} = P_{\mathrm{i,j}} + \mathrm{rand}(M_{\mathrm{i,j}} - I_{\mathrm{i,j}} * P_{\mathrm{i,j}})$

End for

Selection based on new position

If $F_{P1} < F_{Pi}$ then

$P_i = P_{\mathrm{i}}^{P1}$

End if

End for

Roaming

For each walrus i in the population do:

For each dimension j do:

If $F_k < F_i$ then

$\mathrm{v} = x_{\mathrm{i,j}} + \mathrm{rand}(x_{\mathrm{k,j}} - \mathrm{I} * x_{\mathrm{i,j}})$

Else

$x_{\mathrm{i,j}}^{P2} = x_{\mathrm{i,j}} + \mathrm{rand}(x_{\mathrm{i,j}}) - x_{\mathrm{k,j}}$

End if

End for

End for

Escaping

For each walrus i in the population do:

For each dimension j do:

$x_{\mathrm{i,j}}^{P3} = x_{\mathrm{i,j}} + (\mathrm{ul}(ll_{local,\mathrm{j}}^{t}) - \mathrm{rand} * ll_{local,\mathrm{j}}^{t})$

End for

End for

Update the best candidate solution

Identify the strongest walrus (best candidate) based on updated positions and fitness

End for

Step 3: Output the optimal CH nodes

Algorithm 1 CH Selection Using WO

3.3.3. Return the Best Candidate Solutions as the Selected CH Nodes

The process begins with the initialization of the population, where each walrus represents a candidate CH node. The fitness of each candidate is measured using a predefined fitness function that considers node speed, distance from member nodes, and energy consumption. The strongest walrus, representing the best candidate CH is identified based on the fitness values.

3.4. Hybrid Method

To achieve a higher accuracy, a hybrid CNN and LSTM model is proposed for water quality prediction which is shown in Figure 3.

3.4.1. CNN Model

The CNN method is used to extract spatial features from the input data that has multiple convolutional layers followed by pooling layers. The input layer receives the water quality time-series data. These layers apply convolution operations using multiple filters to extract spatial data. The pooling layers are used to minimise the feature map's dimensionality.

3.4.2. LSTM

LSTM network is used to implement the efficient and highly predictive model. This model is used to overcome the limitation of RNN. It is mainly focused on overcoming the issues of vanishing gradient and short contexts. While performing a backpropagation in the LSTM, the neuron gradient value tends to zero. So, there is no need for a neuron propagation which is called an exploding gradient issue.

The LSTM can vanish gradient issues by considering long-term memory dependencies effectively which is given in Figure 4. It has input gates, hidden or forget gates and output gates. The input gate is used to collect data; the hidden gate is used to make a decision which is an intermediate gate between input and output. Next, the output gate is used to provide a valid solution to the network. In this network, the forget gate operation is the main function. It is used to spread

RESEARCH ARTICLE

valid data to the network to avoid irrelevant data. The FORGET Gate is effective in decision-making and can be forwarded to the data by estimating or eliminating cell state for the next step. The data in the forget gate are propagated through the sigmoid activation function. When the output data is not equal to zero, it is forward to evaluate the cell state. If the output data of the forget gate is zero, then the data will be eliminated.

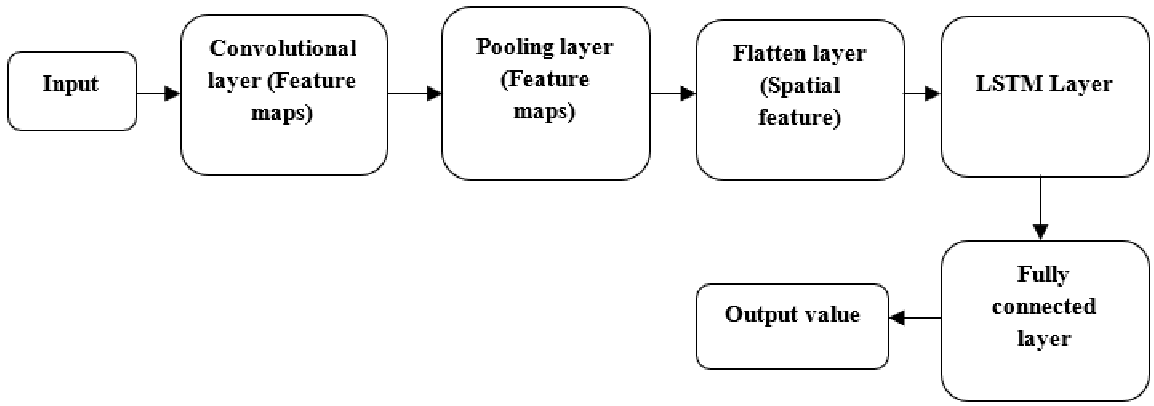

Figure 3 Hybrid CNN+LSTM Model

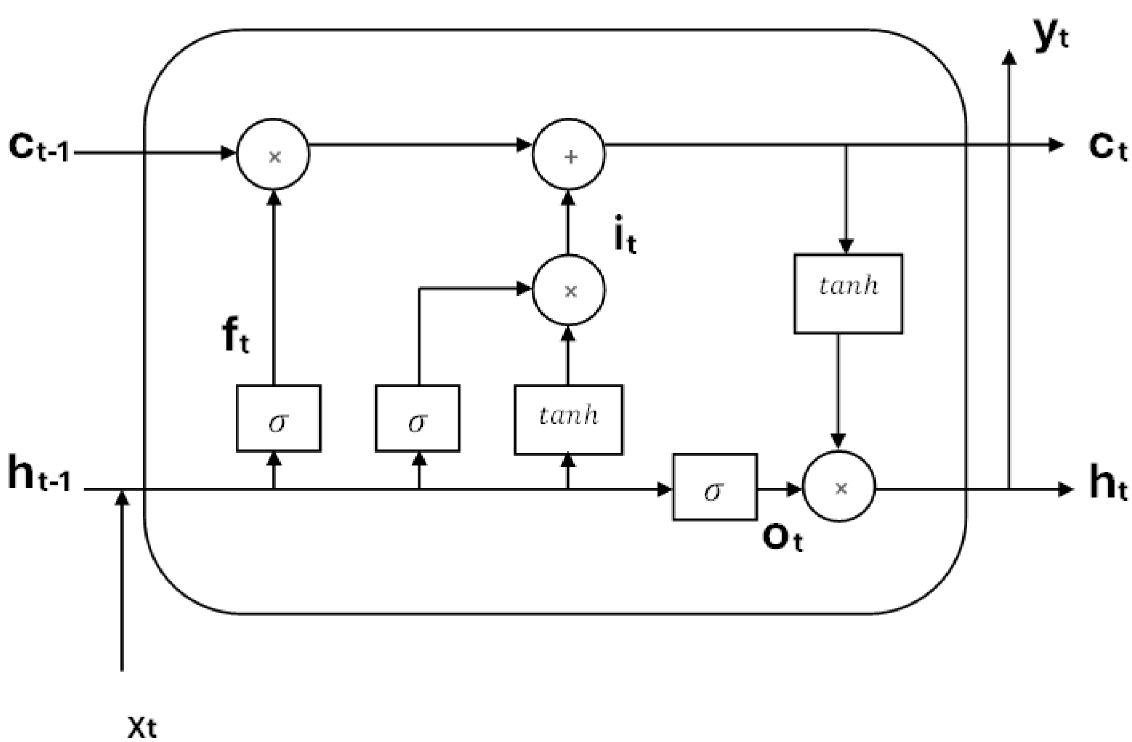

Figure 4 LSTM Model for Prediction

The Input Gate in the LSTM network is used to generate a cell state and validate the data importance. Then the data is sent to the forget gate. This forgets gate can process a data elimination or propagation. The input gate provides a data priority to determine and store the more appropriate data in memory. This gate can be used to perform a tanh activation function to control the network's flow. The relevant data updates a Cell state. Then the forget gate is multiplied with an output gate to avoid the zero values possibilities. Then, the addition of pointwise is done with an input gate to generate the cell state update. Finally, in the LSTM network, the Output gate is helped to choose the next hidden cell state. Then the updated cell state is shifted to the tanh' function. Then this function value is multiplied with the sigmoid response. This response is to select the hidden state of data to be revealed.

The functions of gate outputs and LSTM outputs can be seen below (Eq. 9 – Eq. 16):

$$f_t = \sigma(W_{xf}.x_t + .W_{hf}.h_t - 1 + b_f) \quad (9)$$

$$i_t = \sigma(W_{xi}.x_t + .W_{hi}.h_t - 1 + b_i) \quad (10)$$

$$\bar{c}_t = tanh(W_{xc}.x_t + .W_{hc}.h_t - 1 + b_c) \quad (11)$$

$$c_t = f_t * c_{t-1} + i_t * \bar{c}_t \quad (12)$$

$$O_t = \sigma(W_{xo}.x_t + W_{ho}.h_t - 1 + b_o) \quad (13)$$

$$h_t = O_t * tanh \quad (14)$$

$$y_t = \sigma(W_{hy}.h_t + b_y) \quad (15)$$

$$\sigma(x) = \frac{1}{1+e^{-x}} \quad (16)$$

Where f_t, i_t, O_t represents forward, input and output gate functions with weights of W_{xf}, W_{xi} and W_{xo} and b denotes the bias values respectively.

4. RESULT AND DISCUSSION

In this section, the simulation results are described for the proposed clustering and prediction model using MATLAB software. The simulation parameters are listed in Table 2**.** In the network, the Nodes are deployed and located in the area of 500*500m randomly. The node's quantity is varied with a difference of 20 up to 100 for instance 20,40,60,80 and 100 respectively. The initial node energy is assigned as 10.5J with a 75m transmission range. The value of 0.660W and 0.395W is the power range of transmitting and receiving nodes and also the size of every packet is fitted to 512. The performance metric of this model is considered for an evaluation in terms of delay, energy, PDR and data accuracy respectively.

Table 2 Simulation Parameters Settings

Parameter	Value
Area	500m x 500m
Number of Nodes	20, 40, 60, 80, 100
Initial Node Energy	10.5J
Transmission Range	75m
Transmission Power	0.660W
Receiving Power	0.395W
Packet Size	512 bytes

4.1. Delay Analysis

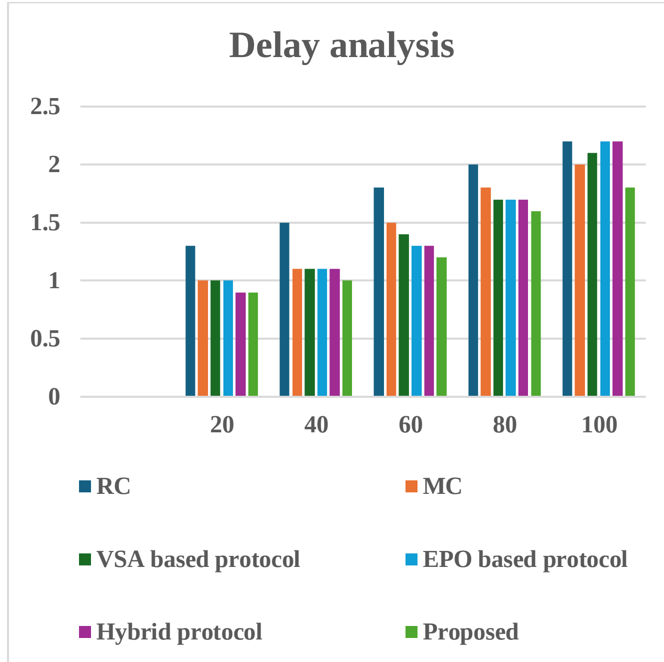

Figure 5 Delay vs Number of Nodes

For a fair comparison, the proposed clustering is compared with Random Clustering (RC), Multi-parameter Clustering (MC), VSA [35], EPO [36] and hybrid algorithms [34]. From the above Figure 5, the x-axis denotes a number of nodes in terms of 20,40,60,80 and 100.

The Y-axis represents the delay in seconds where the time taken to transmit the packets to the source and destination node. When the nodes are increased then the delay of end-to-end network is also increased. As a result, the proposed model delay is compared to previously proposed models.

The output of the graph (Figure 5) shows that the proposed model has obtained a very minimum time delay for packet transmission than the VSA, EPO and Hybrid methods.

4.2. PDR Analysis

From the Figure 6, the x-axis represents a node in terms of 20,40,60,80 and 100. The Y-axis indicates PDR which is the quality ratio of a packet to the source and destination node with a minimum error. When the node's count is increased then the PDR is also increased. As a result, the proposed model is compared with an existing method.

The output of the graph shows that the proposed model has obtained a higher PDR value than other methods. Therefore, the performance of the proposed method is observed that the PDR value is 26% greater than the VSA, EPO and Hybrid methods.

RESEARCH ARTICLE

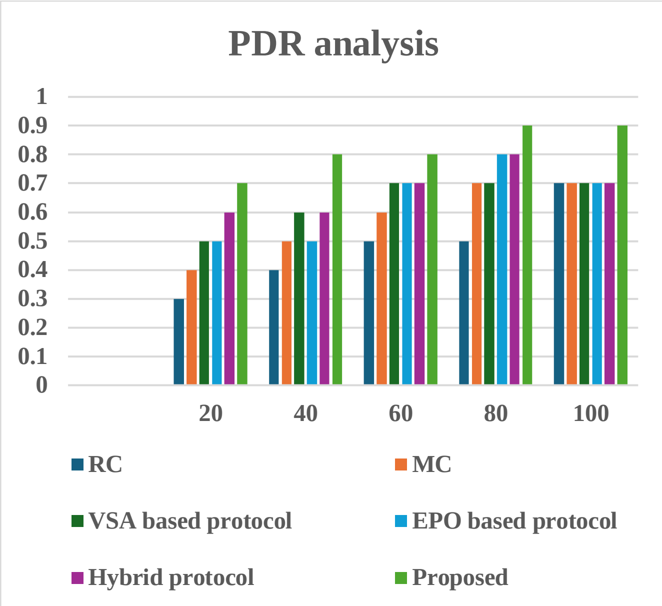

Figure 6 PDR vs Number of Nodes

4.3. Energy Analysis

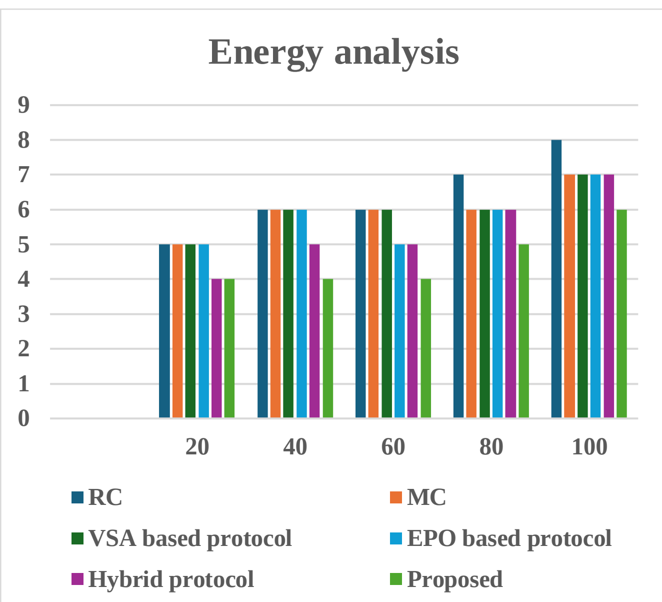

Figure 7 Energy Consumption vs Number of Nodes

From the Figure 7, the x-axis denotes a number of nodes in terms of 20,40,60,80 and 100. The Y-axis represents the energy consumption in Joule which is the overall energy consumed for an end-to-end packet delivery. When the number of nodes is increased then the energy consumption of the end-to-end network is also increased. As a result, the proposed model's energy consumption is compared with other models. The output of the graph shows that the proposed model has consumed a much less amount of energy than other methods.

4.4. Lifetime Analysis

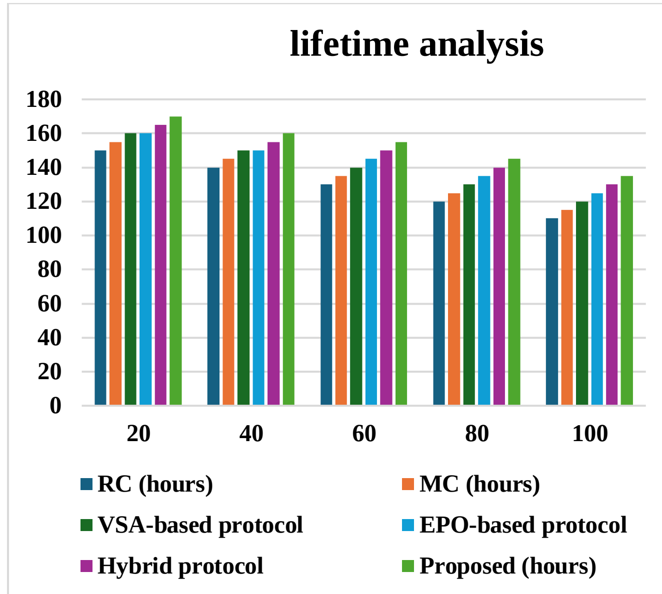

Figure 8 Lifetime vs Number of Nodes

From the Figure 8, the x-axis represents a node in terms of 20,40,60,80 and 100. The Y-axis represents the lifetime of the nodes in hours. When the number of nodes is increased then the lifetime of the network is reduced. The output of the graph shows that the proposed model has a better lifetime than other methods due to the optimal selection of cluster heads.

4.5. Throughput Analysis

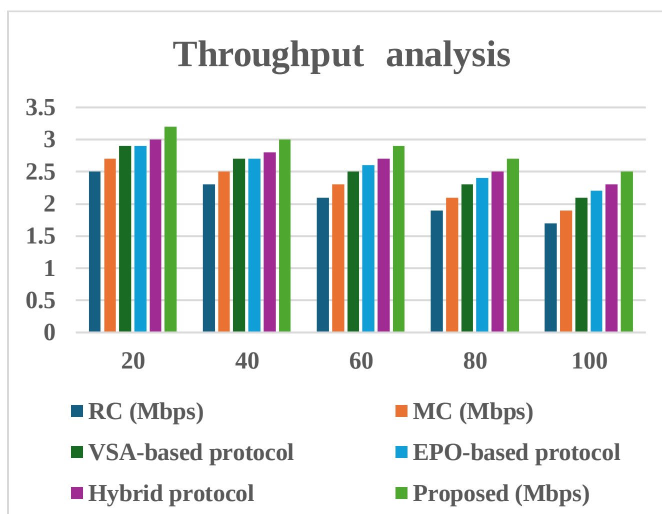

Figure 9 Throughput vs Number of Nodes

RESEARCH ARTICLE

From the Figure 9, the x-axis denotes a number of nodes in terms of 20,40,60,80 and 100. The Y-axis represents the throughput of the communication. The output of the graph shows that the proposed model has better results than others even when increasing network size.

4.6. Accuracy Analysis

The water quality data is collected from the Kaggle website (https://www.kaggle.com/datasets/adityakadiwal/water-potability). The data set includes different water quality parameters. The dataset includes a total of 3277 samples. The visualization of a data set is given in Figure 10. From the data set, 80% data is trained by a hybrid model and the remaining 20% data is used for validation purposes.

The overall performance of the model is given in Table 3. The hybrid model achieved better performance than other models. The model attained a higher accuracy of 97.8%, with a Precision score of 95.2 %, and a Recall score of 85.4.

The confusion matrix plot of the proposed classifier is given in Figure 11. The diagonal cells represent the correct predictions of true positives and true negatives. The off-diagonal cells represent the incorrect predictions of false positives and false negatives. It proves the suitability of the proposed model for water quality prediction in real-time.

	A	B	C	D	E	F	G	H	I	J
1	ph	Hardness	Solids	Chloramin	Sulfate	Conductiv	Organic_c	Trihalomet	Turbidity	Potability
2		204.8905	20791.32	7.300212	368.5164	564.3087	10.37978	86.99097	2.963135	0
3	3.71608	129.4229	18630.06	6.635246		592.8854	15.18001	56.32908	4.500656	0
4	8.099124	224.2363	19909.54	9.275884		418.6062	16.86864	66.42009	3.055934	0
5	8.316766	214.3734	22018.42	8.059332	356.8861	363.2665	18.43652	100.3417	4.628771	0
6	9.092223	181.1015	17978.99	6.5466	310.1357	398.4108	11.55828	31.99799	4.075075	0
7	5.584087	188.3133	28748.69	7.544869	326.6784	280.4679	8.399735	54.91786	2.559708	0
8	10.22386	248.0717	28749.72	7.513408	393.6634	283.6516	13.7897	84.60356	2.672989	0
9	8.635849	203.3615	13672.09	4.563009	303.3098	474.6076	12.36382	62.79831	4.401425	0
10		118.9886	14285.58	7.804174	268.6469	389.3756	12.70605	53.92885	3.595017	0
11	11.18028	227.2315	25484.51	9.0772	404.0416	563.8855	17.92781	71.9766	4.370562	0

Figure 10 Trained Data Set Visualization

Table 3 Performance Analysis

Model	Accuracy	Precision	Recall	F1 Score
Proposed model	97.8	95.2	85.4	89.6
GRU	95.50	92.00	80.00	85.50
LSTM	88.75	84.50	65.00	73.50
CNN	86.50	82.00	60.00	70.00

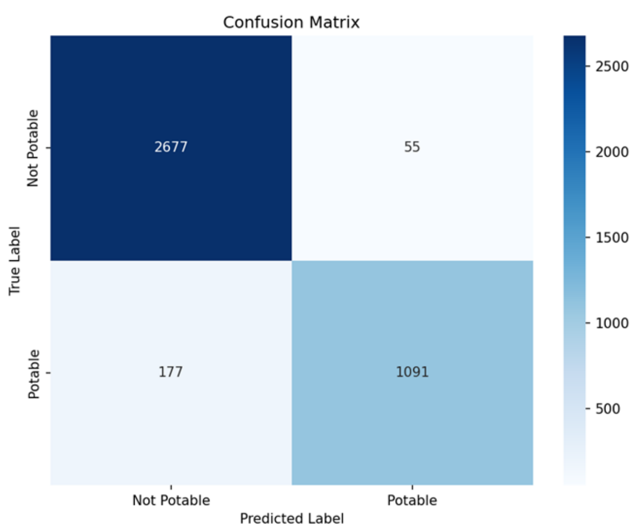

Figure 11Confusion Matrix Plot

RESEARCH ARTICLE

5. CONCLUSION

This work presents a new approach for river water quality monitoring using clustering and deep learning models in MANETs. The use of the WOA for selecting optimal CHs improves network efficiency with reduced energy consumption and delays. The proposed clustering protocol shows better performance in all evaluation parameters. Moreover, the hybrid CNN-LSTM model is used for water quality prediction. It effectively captures complex spatial and temporal patterns in the data. The hybrid model achieved an accuracy of 97.8%, outperforming other models such as GRU, LSTM, and CNN in terms of precision, recall, and F1 score rates. In future work, a federated learning approach will be integrated into the water quality monitoring system using MANETs. This integration will reduce communication costs and further improve network performance.

REFERENCES

[1] Z. Jian et al., "Assessment of Spatial-Temporal Patterns of Surface Water Quality in the Min River (China) and Implications for Management," 2011 International Conference on Computer Distributed Control and Intelligent Environmental Monitoring, Changsha, China, 2011, pp. 1983-1989, doi: 10.1109/CDCIEM.2011.499.

[2] Liming Zhang and Haowen Yan, "Implementation of a GIS-based water quality standards syntaxis and basin water quality prediction system," 2012 International Symposium on Geomatics for Integrated Water Resource Management, Lanzhou, 2012, pp. 1-4, doi: 10.1109/GIWRM.2012.6349656.

[3] V. Dattana, A. Kumar, A. Kush and S. I. Ali Kazmi, "Manet for Stable Data flow in Smart home and Smart city," 2019 4th MEC International Conference on Big Data and Smart City (ICBDSC), Muscat, Oman, 2019, pp. 1-4.

[4] J. Zhang, C. Liu and T. Zheng, "Smart Integrated MANET-DTN Scheme for Network Adaptation Enhancement in Emergency Communication," 2023 IEEE 6th Information Technology,Networking,Electronic and Automation Control Conference (ITNEC), Chongqing, China, 2023, pp. 1591-1594.

[5] A. Katal, M. Wazid, R. S. Sachan, D. P. Singh and R. H. Goudar, "Effective Clustering Technique for Selecting Cluster Heads and Super Cluster Head in MANET," 2013 International Conference on Machine Intelligence and Research Advancement, Katra, India, 2013, pp. 1-6, doi: 10.1109/ICMIRA.2013.8.

[6] Men Baohui, Zhao Meiling, Xu Qingli, Li Rongbo, Ye Xiaoning and Zhang Yancheng, "Water quality assessment of Wenyu River based on attribute recognition method," 2012 International Conference on Computer Science and Information Processing (CSIP), Xi'an, China, 2012, pp. 140-143, doi: 10.1109/CSIP.2012.6308814.

[7] H. P. Hanifah and S. H. Supangkat, "IoT-based River Water Quality Monitoring Design for Smart Environments in Cimahi City," 2019 International Conference on Electrical Engineering and Informatics (ICEEI), Bandung, Indonesia, 2019, pp. 496-499, doi: 10.1109/ICEEI47359.2019.8988883.

[8] R. Hartono, M. A. Safi'Ie, N. M. Yoeseph, S. A. T. Bawono, Hartatik and A. Aziz, "Improved Data Transmission of Smart Water Quality Sensor Devices in Bengawan Solo River with LoRa," 2022 1st International Conference on Smart Technology, Applied Informatics, and Engineering (APICS), Surakarta, Indonesia, 2022, pp. 171-174, doi: 10.1109/APICS56469.2022.9918709.

[9] A.P. Kogekar, R. Nayak and U. C. Pati, "A CNN-GRU-SVR based Deep Hybrid Model for Water Quality Forecasting of the River Ganga," 2021 International Conference on Artificial Intelligence and Machine Vision (AIMV), Gandhinagar, India, 2021, pp. 1-6, doi: 10.1109/AIMV53313.2021.9670916.

[10] L. Arya and G. Srivastava, "Fuzzy Models for Water Quality Assessment," 2019 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT), Ghaziabad, India, 2019, pp. 1-6, doi: 10.1109/ICICT46931.2019.8977693.

[11] L. Jia, N. Yen and Y. Pei, "Spatial and Temporal Water Quality Data Prediction of Transboundary Watershed Using Multiview Neural Network Coupling," in IEEE Transactions on Geoscience and Remote Sensing, vol. 61, pp. 1-16, 2023, Art no. 3000816, doi: 10.1109/TGRS.2023.3334291.

[12] Wang Jingmeng, Guo Xiaoyu, Zhao Wenji and Meng Xiangang, "Research on water environmental quality evaluation and characteristics analysis of TongHui River," 2011 International Symposium on Water Resource and Environmental Protection, Xi'an, China, 2011, pp. 1066-1069, doi: 10.1109/ISWREP.2011.5893198.

[13] Yunchao Jiang and Zhongren Nan, "Assessment of river water quality using uncertainly mathematical model: A case Study of Yellow River, China," 2012 International Symposium on Geomatics for Integrated Water Resource Management, Lanzhou, China, 2012, pp. 1-4, doi: 10.1109/GIWRM.2012.6349563.

[14] Z. -l. Hao, Y. -y. Zhang and M. -q. Feng, "Water quality assessment based on BP network and its application," 2011 International Symposium on Water Resource and Environmental Protection, Xi'an, China, 2011, pp. 872-876, doi: 10.1109/ISWREP.2011.5893150.

[15] Kalphana, K. R., S. Aanjankumar, M. Surya, M. S. Ramadevi, K. R. Ramela, T. Anitha, N. Nagaprasad, and Ramaswamy Krishnaraj. "Prediction of android ransomware with deep learning model using hybrid cryptography." Scientific Reports 14, no. 1 (2024): 22351.

[16] A.K. Shukla, C. S. P. Ojha and R. D. Garg, "Surface water quality assessment of Ganga River Basin, India using index mapping," 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 2017, pp. 5609-5612, doi: 10.1109/IGARSS.2017.8128277.

[17] Kumar, S. Aanjan, P. Karthikeyan, S. Aanjana Devi, S. Poonkuntran, V. Palanisamy, and V. Navatharani. "Protecting Medical Images Using Deep Learning Fuzzy Extractor Model." In Deep Learning for Smart Healthcare, pp. 183-203, 2024.

[18] R. P. N. Budiarti, A. Tjahjono, M. Hariadi and M. H. Purnomo, "Development of IoT for Automated Water Quality Monitoring System," 2019 International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE), Jember, Indonesia, 2019, pp. 211-216, doi: 10.1109/ICOMITEE.2019.8920900.

[19] A.L. Lopez, N. A. Haripriya, K. Raveendran, S. Baby and C. V. Priya, "Water quality prediction system using LSTM NN and IoT," 2021 IEEE International Power and Renewable Energy Conference (IPRECON), Kollam, India, 2021, pp. 1-6, doi: 10.1109/IPRECON52453.2021.9640938.

[20] Kumar, S. Aanjan, Monoj Kumar Muchahari, S. Poonkuntran, L. Sathish Kumar, Rajesh Kumar Dhanaraj, and P. Karthikeyan. "Application of hybrid capsule network model for malaria parasite detection on microscopic blood smear images." Multimedia Tools and Applications (2024): 1-27.

[21] Chellaswamy C and K. K, "Smart River Water Quality and Level Monitoring: a Hybrid Neural Network Approach," 2023 International Conference on Advances in Intelligent Computing and Applications (AICAPS), Kochi, India, 2023, pp. 1-6, doi: 10.1109/AICAPS57044.2023.10074495.

[22] Q. Ye, X. Yang, C. Chen and J. Wang, "River Water Quality Parameters Prediction Method Based on LSTM-RNN Model," 2019 Chinese Control And Decision Conference (CCDC), Nanchang, China, 2019, pp. 3024-3028, doi: 10.1109/CCDC.2019.883288.

[23] A. A. Nayan, M. G. Kibria, M. O. Rahman and J. Saha, "River Water Quality Analysis and Prediction Using GBM," 2020 2nd International Conference on Advanced Information and Communication Technology

RESEARCH ARTICLE

(ICAICT), Dhaka, Bangladesh, 2020, pp. 219-224, doi: 10.1109/ICAICT51780.2020.9333492.

[24] A.Dengfeng, J. Qichun and Z. Hongcai, "Fuzzy Synthetic Evaluation of Water Quality of Hei River System," 2011 Third International Conference on Measuring Technology and Mechatronics Automation, Shanghai, China, 2011, pp. 280-283, doi: 10.1109/ICMTMA.2011.357.

[25] Aanjanadevi, S., S. Aanjankumar, K. R. Ramela, and V. Palanisamy. "Face Attribute Convolutional Neural Network System for Data Security with Improved Crypto Biometrics." Computer Systems Science & Engineering 46, no. 1 ,2023.

[26] L. Guo and D. Fu, "River Water Quality Prediction Model Based on PCA-APSO-ELM Neural Network," 2023 6th International Conference on Artificial Intelligence and Big Data (ICAIBD), Chengdu, China, 2023, pp. 512-517, doi: 10.1109/ICAIBD57115.2023.10206249.

[27] D. Zhang, R. Chang, H. Wang, Y. Wang, H. Wang and S. Chen, "Predicting Water Quality Based On EEMD And LSTM Networks," 2021 33rd Chinese Control and Decision Conference (CCDC), Kunming, China, 2021, pp. 2372-2377, doi: 10.1109/CCDC52312.2021.9602800.

[28] Z. Guo, R. Gai, S. Qin and P. Wang, "CNN-BiLSTM for water quality prediction method based on attention mechanism," 2023 IEEE Smart World Congress (SWC), Portsmouth, United Kingdom, 2023, pp. 1-6, doi: 10.1109/SWC57546.2023.10448856.

[29] A.P. Kogekar, R. Nayak and U. C. Pati, "Forecasting of Water Quality for the River Ganga using Univariate Time-series Models," 2021 8th International Conference on Smart Computing and Communications (ICSCC), Kochi, Kerala, India, 2021, pp. 52-57, doi: 10.1109/ICSCC51209.2021.9528216.

[30] H. Sun and Y. He, "Research and Application of Water Quality Evaluation of a Certain Section of Yangtze River Based on Fuzzy Neural Network," 2017 International Conference on Industrial Informatics - Computing Technology, Intelligent Technology, Industrial Information Integration (ICIICII), Wuhan, China, 2017, pp. 301-304, doi: 10.1109/ICIICII.2017.39.

[31] P. Siagian, A. Sagala, S. Hutauruk and G. F. Panggabean, "Design of WSNs Sensor for River Water Quality Monitoring System for Neolissochillus thienemanni sumateranus Lifes," 2022 IEEE International Conference of Computer Science and Information Technology (ICOSNIKOM), Laguboti, North Sumatra, Indonesia, 2022, pp. 1-6

[32] Z. Song, C. Zhang and Z. Jin, "Water Quality Prediction in Zhengzhou City Based on ESN Model," 2024 IEEE 3rd International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA), Changchun, China, 2024, pp. 1211-1215, doi: 10.1109/EEBDA60612.2024.10485769.

[33] S. Chopade, H. P. Gupta, R. Mishra, A. Oswal, P. Kumari and T. Dutta, "A Sensors-Based River Water Quality Assessment System Using Deep Neural Network," in IEEE Internet of Things Journal, vol. 9, no. 16, pp. 14375-14384, 15 Aug.15, 2022, doi: 10.1109/JIOT.2021.3078892.

[34] A., Rajaram, & A., Baskar. (2023). Hybrid Optimization-Based Multi-Path Routing for Dynamic Cluster-Based MANET. Cybernetics and Systems, 1–23. https://doi.org/10.1080/01969722.2023.2166249.

[35] Venkatasubramanian, S. (2023). Optimal Cluster Head Selection Using Vortex Search Algorithm with Deep Learning-Based Multipath Routing in MANET. In: Joby, P.P., Balas, V.E., Palanisamy, R. (eds) IoT Based Control Networks and Intelligent Systems. Lecture Notes in Networks and Systems, vol 528. Springer.

[36] Hamza, F., Vigila, S.M.C. (2023). An Energy-Efficient Cluster Head Selection in MANETs Using Emperor Penguin Optimization Fuzzy Genetic Algorithm. In: Mahapatra, R.P., Peddoju, S.K., Roy, S., Parwekar, P. (eds) Proceedings of International Conference on Recent Trends in Computing. Lecture Notes in Networks and Systems, vol 600. Springer.

[37] Kavitha, V. P., Sakthivel, B., Deivasigamani, S., Jayaram, K., Badlishah Ahmad, R., & Sory Keita, I. (2024). An Efficient Modified Black Widow Optimized Node Localization in Wireless Sensor Network. IETE Journal of Research, 1–10. https://doi.org/10.1080/03772063.2024.2384486.

[38] Trojovský, P., Dehghani, M. A new bio-inspired metaheuristic algorithm for solving optimization problems based on walruses behavior. Sci Rep 13, 8775 (2023). https://doi.org/10.1038/s41598-023-35863-5.

Authors

Ms. D. Manopriya, MCA., from Vivekananda Business School for women, Tiruchengode. She also completed MPhil (Computer Science)., in Sri Vasavi College, Erode. She is Research Scholar in Sri Krishna Adithya Arts and Science College, Coimbatore. She has published more papers , journals & chapters in various topics and also attended many conferences & workshops at National & International level. She has an academic experience of 4 years 10 Months.

Dr. M. Manimaran is a Doctorate from Anna University and M.E in Computer Science & Engineering from SVS College of Engineering & Master of Computer Applications in Amrita Vishwa Vidyapeetham, Ettimadai, Coimbatore. He also completed his M.Phil in Computer Science in Bharathiar University. He has an academic experience of 17 years and has worked at various levels up to Principal & Associate Professor. He has 8 publications, in peer reviewed International/National journals with high impact factor and has 24 publications in various International Conferences held in India and abroad. Dr.M.Manimaran has been an expert speaker in various Faculty Development Programmes and has organized a number of Conferences, Workshops and Short term courses. He has authored a book titled Computer Networks. He was a LIfe Time Membership in International Association of Engineers (IAENG). He was awarded a Best Eduprenaur in the Year 2018 from Nehru Group of Institutions, Coimbatore, Tamil Nadu.He was also awarded a global peace award by Miss Peace Ambassador Global for his outstanding achievements in field of Social Development and Education in the year 2022 at New Delhi.

How to cite this article:

D. Manopriya, M. Manimaran, “Integrating MANETs and Hybrid Deep Learning for Enhanced River Water Quality Monitoring”, International Journal of Computer Networks and Applications (IJCNA), 11(5), PP: 660-672, 2024, DOI: 10.22247/ijcna/2024/41.