International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-2, July 2019

Optimal Features-based Channel Selection and Neural Network Learning for LTE applications

Divya Mohan, Geetha Mary. A

***Abstract*: *Nowadays, the number of internet users is increasing vastly. Hence, predicting the number of channels for user's communication is a major task. Therefore, scheduling of all traffic flows within the communication services in Long Term Evolution (LTE) scheduling is done by verifying the channel information and user availability in the network. In view of that, this paper proposes a novel feature extraction and classification method to evaluate the user availability status and the channel information for the betterment of communication within the LTE network. For preprocessing, we present a Fast Independent Component Analysis (FICA) method which incorporates dimensionality reduction in the given input feature data during the training course. In this work, feature extraction algorithm is used to extract the network feature and its degree of angle by means of distance between the features set. Subsequently, analyzes the extracted categories of network architecture by the analyzing weight value of that attribute based on the responses. As a result both the feature identification and framing of network structure increases the performance of data mining analysis. Hence, proposing a novel optimization technique like Particle Connected Cuckoo Search (PCCS) optimization algorithm for selecting the best features in the training feature set. Based on the extracted features, the classification method is performed in order to predict the network category by using Multilevel Neural Network (MLNN) classification technique. At this point, a novel kernel model for classification process is incorporated to reduce the time complexity. After that, the information is passed to the LTE scheduling system for providing enhanced communication. The comparative analysis between proposed techniques with the existing methods such as naive Bayes, stacking C, DTNB, random forest, J48, Ridor, decision table, zero R, grading in terms of accuracy, RMSE, CCI, ICI assures that the effectiveness of the proposed MLNN classification method.*

Index Terms: Classification, Component Analysis, Data mining, Fast Independent, Feature Extraction, LTE.

I. INTRODUCTION

The fast-growing demand for network service is the major part of future generation communication network. The major issues in the communication networks deals with user requirements. For satisfying the user needs, the selection of optimal packets for scheduling process of radio access and networks through Long Term Evolution (LTE) is done.

Revised Manuscript Received on 30 July 2019.
***** Correspondence Author
Divya Mohan*, School of Computer science and Engineering , Vellore Institute of technology , Vellore, India.
Geetha Mary.A , School of Computer science and Engineering , Vellore Institute of technology , Vellore, India

The Universal Mobile Telecommunication Systems (UMTS) is anticipated by LTE scheme for providing good spectrum utilization,
great flexibility, low latency and user data rates which needs to be higher in both uplink and downlink direction to the cellular network users. Thus, the increasing demand in mobile data traffic is considered to be the significant interest of operators to optimize the resource allocation process [7].
The mobile-based applications is based on the relevant resource allocation which in turn depends on two categories as follows: opportunities scheduling and fair proportional scheduling. The extraction of an optimal solution for application-aware resource scheduling problem in LTE systems needs further discussion. In recent days, the promising multiple access (MA) technique is implemented on the recommendation of non-orthogonal multiple access (NOMA) [5]. The efficiency of mobile network spectral communication is enhanced by use of this technique. In NOMA technique, some bandwidth resource provides an issue in the sub-carrier, poor channel selection and lower spectral efficiency. Therefore, the use of NOMA technique enables each user to access all subcarrier channels with strong channel condition and improves the spectral efficiency. Hence, the system is made to accomplish the demand requirements of 5G such as ultra-low latency, ultra-high connectivity by which the user utilizes different channel conditions in a correct manner.

The LTE/LTE-Advanced and WiMAX are introduced as higher data rate wireless access for increasing the demand in mobile users [20].The channel selection is performed based on the evaluation process within the LTE architecture majorly in evolving node (EN) or access points (AP) in the centralized network topology. Hence, the large number of communicating channels gets congested in the centralized network topology. As a result the communication service, selects the best optimal channel in the network. The device to device communication includes some limitations, such as, interference occurrence, low spectral efficiency and low system capacity. Therefore, the channel intrusion is identified by suitable technique in order to alleviate these problems [4]. The acknowledgement of channel bonding variations, adjacent channel interference (ACI), access time, and the corresponding power consumption for transmission were derived in this system which is the major aspects to decrease the bandwidth capacity. Here the interpretation of channel bonding is made through the set of contiguous non-overlapping wireless channels which supports an efficient performance in order to provide higher bandwidth. Within a small duration, the collaborative beam (CB) of information is transmitted to large chunks of data. In communication-based networks, and channel selection are the promising approach for providing dynamic spectrum access (DSA).

The major benefits of channel selection is to provide less complexity, higher bandwidth, and larger channel capacity. The wireless sensor network based DSA recommends relevant channel selection based on two categories such as licensed and unlicensed bands. The major challenges of CB in CRNs are addressed in future in turn to resolve the problems. The path discovery, movement behavior analysis, and location prediction are learned from various applications of trajectory data mining [8].The moving objects generates the trajectory data and it is collected from multiple data sources.

The preprocessing, confidentiality, security, inquiry processing, data management, and trajectory data mining are the five different phases of the trajectory mining process. In this section, the brief review of approaches is discussed such as data complexity; low channel capacity, energy degradation and low spectral efficiency which are considered as the major issues to be solved for a better recommendation. This paper proposes the novel algorithms to alleviate the issues in existing recommendation approach.

The technical contributions of proposed method are listed as follows:

1) Fast independent component analysis (FICA) method is utilized to reduce the redundant features which are present in the input or data attributes.

2) In training, the best features set is selected by particle connected cuckoo search (PCCS) optimization algorithm.

3) The proposed multilevel neural network (MLNN) classification method is used in order to predict the label of network category based upon their classification. The paper is organized as follows: The detailed description of the related works of the methods involved in the relevant recommendation of services is discussed in related work section. The implementation process of multilevel neural network (MLNN) classification method is described in third section. The performance analysis and the evaluation of the proposed system based on some performance measures is discuss in section IV .Section V deals with the comparative analysis of the proposed approach with that of the existing methods is conferred. Finally, the conclusion about the proposed approach is presented in section VI.

II. RELATED WORK

This section discusses the related work based on the various recommendation models which is available traditionally. The evolution of feature extraction and classification method in the data mining to retrieve the relevant information of the network topology for scheduling the communication network.[19] compared the different packet scheduling algorithm for 3GPP UTRAN long term evolution (LTE). They evaluated the performance of packet scheduling (PS) and opposed the resources due to the existence of different types of traffic with different Quality of Service (QoS) requirements. The throughput fairness between the users was controlled and utilized by time and frequency domain packets. The comparative results of improved fairness and high-quality cell throughput was obtained.[11] Suggested the LTE scheduling downlink channel policies with time and frequency domain. They promoted several flourishing scheduling algorithms with the different purpose and allowed wide discretion of resource allocation. They have conducted the experimental analysis with different scheduling algorithm such as saturated UDP and TCP traffic sources. The relationship between the flat and frequency selective channels and the schedules of time and frequency domains were considered. The major issue found in this system is low connectivity services. [17] Aimed to examine the large-scale coordination scheduling algorithm applicable in LTE-Advanced networks. The large scale coordination obtained from the layered solution such as the clustering of few cells also coordinated the clustering of coordinated calls. They developed the model of small-scale coordination and large-scale coordination since it reduced the optimization problem. The optimality problem between the networks was a major issue, hence two efficient heuristic techniques were recommended for solving the issue which was expensive. Hence, the coordination of scheduling algorithm was the major problem that limits the efficiency of the system as it required some energy to perform. [21] Here it was planned that the LTE networks gets adapted to the framework for pricing aware resource scheduling (PARS). This framework aimed to conquer the source scheduling performance and increased operator revenue contradictory. They established three layer schedules in scheduling module to assign resources based on the traffic amount, packet delay, channel condition, and user level. The users limited the amount of money charged in pricing module that modified the demand of price elasticity. The high cost of mobile devices for charging and collecting was the major problem , which in turn reduced the overall performance. [9] Predicted the best downlink resource distribution based on the greedy-knapsack algorithm in LTE networks. The system performance maximized due to an optimal selection set of users without exceeding the available bandwidth capacity. The relationship between the QoS and throughput described for each application by class-based ranking functions. The overload traffic was the major problem as it causes resource allocation problems also considered to be computationally expensive scheme. [14] Developed a unified graph labelling algorithm in order to alleviate the challenges such as channel utility, power consumption and the minimum number of channels. They utilized SC-FDMA channel allocation to find an optimal path in an acyclic graph. Thus, the global optimality guarantee of the algorithm was justified from the two specific input scenarios such as user-invariant and channel-invariant with the strong polynomial-time complexity. The channel power utilization is the major issue in the recommended model.[3] Recommended for enhancing the effects of FWM that utilized the effective bandwidth channel allocation scheme. They utilized to calculate the effects of four-wave mixing (FWM) parameters in the (UDWDM-PON) system that abbreviated from the ultra-dense wavelength division multiplexing passive optical network. The creation of several blocks of the grouped signal channels with different channel spacing. The recommended scheme completely did not reduce the FWM effects , which was a major drawback.[10] Exploited an OFDMA-based wireless network that allowed for the green device to device communication. They utilized the effective algorithm to solve the mode selection, power assignment and channel allocation in a polynomial time and reduced power consumption due to extend the user required data rates. The reliable communication done by mode selection from the cellular networks with low energy efficiency and high energy cost was a major issue.[8] utilized the full duplex cognitive femtocell network (FDCFN) by ensuring the users with the quality of service (QoS) in the form of signal to interference plus noise ratios (SINR).

They developed the iterative framework to solve the problem of power control, channel allocation of the mixed integer nonlinear problem (MINLP) hence selected joint duplex mode as a solution. The distributed power control in a full duplex was the major problem with the recommendation technique. [15] Referred the ranking channel resource distribution scheme was applied semi-Markov decision policy (SMDP) scheme.

In cognitive enabled vehicular ad hoc network environment (VANET) has a shortage problem in channel allocation which was solved by using SMDP policy. The design for ranking was measured using two services such as primary users (PUs) and secondary users (SUs). The ambiguity between the increased demand in the diverse vehicular application and lack of spectrum resource was the main issues with this model.

[21] Offered DXD scheme that appreciated from Dynamic Scheduling with Extensible Allocation and Dispersed Offsets. This scheme revealed two algorithms such as DRX parameters decision and DRX aware scheduling. The relationship between the QoS constraints and channel condition were used to determine the DRX period, and the DRX parameters decision and the second algorithm were used to determine the extended duration due to QoS which will not be affected by DRX. The major issues of power consumption reduced only limited range. [18] Suggested LTE downlink cellular network that's designed and evaluated the performance of scheduling algorithm. Then adapted Round Robin, Max Rate, and Proportional Fairness scheduling algorithms to model and evaluated the performance. The scheduling algorithm allocated the radio resource to users was the major requirement to reduce the throughput and QoS. [2] Planned to recover the video traffic communication signals that have to be removed based on novel delay based scheduling algorithm in LTE network. The new scheduling algorithm accomplished with the delay that increases throughput performance of the system. The lack of video traffic related to the grouping of products or users yielded the poor recommendation performance.[1] referred LTE real-time traffic was utilized in the exponential rule based the algorithm. They evaluated the performance of packet scheduling schemes. They utilized four different stages of the recommended scheme. They have simplified the existing exponential rule (sEXP rule), improved EXP rule, maximum throughput with the EXP rule, and improved EXP rule with maximum throughput. The system capacity of the recommendation system has highly affected that dependent on limited radio resources. [5] Exploited the novel downlink scheduling system focused on the traffic on the communication system. They utilized the TD-LTE technology in various information flows of eNodeB that produced two classification mappings. Primarily, correlation of important weight factor with that of flow packet information is carried out in order to obtain relevant traffic safety information. Secondarily, connecting the service type importance weight with that of the traffic flow packet information is related to QoS requirements. Thus, scheduled the overall order is done based on the connection. The overall flow which was produced was of low quality and considered to be the major problem as it reduces the throughput and user fairness. The sensitivity, specificity, accuracy, F-measure is considered as the major parameters for achieving best performance.

III. OPTIMAL FEATURE-BASED CHANNEL SELECTION AND NEURAL NETWORK LEARNING FOR LTE APPLICATIONS

This section discusses the implementation details of proposed multivariate learning of neural network classification model for the recommendation of communication. Fig .1 shows the workflow of MLNN to predict the best channel selection and scheduling. In LTE network architecture, the system can start the scheduling for communication process. This block extracts the relevant network parameters from the recommended

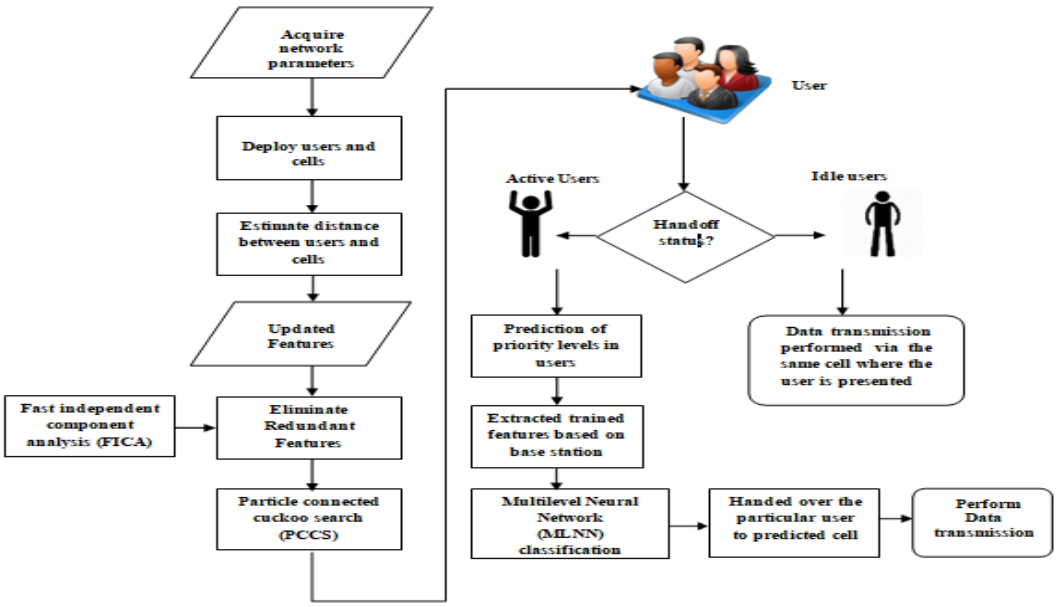


Fig 1. Workflow of proposed MLNN

system and it also requires the input of extracted user details and neighbourhood verification. Based on the parameters, extract the interrelated user details.

Initially, all network parameters are required to deploy the total number of users and corresponding neighbourhood cells (received signals from the nearer tower). The users and neighbourhood cell estimates the distance and updates their required features.

In preprocessing, the Fast Independent Component Analysis (FICA) method is considered as it highly reduces the dimensionality and time complexity. After that, non-redundant features are fed to Particle Connected Cuckoo Search (PCCS) optimization algorithm. This PCCS algorithm is used to select the best features from the feature set effectively.

It checks the available user handoff status which is active or idle. If the receiving hand-off status is prevailing with idle users, then the data transmission is not performed. If the handoff status is not idle, then perform the active user's process. The objective function is performed within the active users block to evaluate the attribute, user status and utilized bandwidth status, which is then transferred to several nodes with the aim of maintaining good bandwidth availability, mobility and trust assess. This information is necessary to prioritize the user status and to get trained. These trained features are the best-optimized results of the PCCS algorithm. If the optimized feature is used to select the feature index value, then the index values are employed to classify the category by Multilevel Neural Network (MLNN) classification method. The prediction and labelling of network category which is assigned to each cell with users by the help of given network information is passed to LTE scheduling system for effective data transmission. The major process in the proposed work is to reduce the time consumption. The following stages are listed below,

- Preprocessing
- Feature Extraction
- Optimization
- Classification

The detailed description of each process in proposed work is presented in next sub-sections. Table I present the variables or user attributes used in the proposed algorithm.

Table I. Symbol and Descriptions

List Of Variables	Description
W_i	Weight of current channel
$W_i^{increase}$	Weight of update channel
X	Input channel information
$\varphi(.)$	Non-quadratic function
W_a	Weight of neighboring channel
Strfind	Status of the user details (enable or disable)
UE_{id}	User equipment id
σ and β	Levy's flights co-efficient
Pa	Probability function of the nest
K	Randomly select the nest (ex: 1, 2, 3…)

A. Pre-Processing

In preprocessing, the unwanted data is reduced to provide the accurate the results. Here, the novel fast Independent Component Analysis (FICA) method is utilized to reduce the redundant data. In preprocessing the network parameters are highly reduced in order to reduce dimensionality issues in the given network data within the trained data. This is done with the help of a novel Fast Independent Component Analysis (FICA) method. FICA is a computational and statistical analysis method that shows the unidentified factors. It describes important irregularities and uses de-correlation of higher order statistical analysis. The input of FICA model is the linear coefficient, which is mutually independent of each other. The algorithm to remove the unwanted features from the network parameters is listed as follows.

Algorithm I: Fast Independent Component Analysis

Input: Network parameters
Output: Maximum Weight value of the features from the reviews

For each network parameter review
// i= 1:n, n- Number of channels

End For;
Step 1: Initialize W_i for first, second, current in random

Step 2: If i= 1

Step 3:$W_i^{increase} = E\left(\varphi'(W_i^T X)\right)W_i - E\left(X\varphi(W_i^T X)\right)$
Step 4:$W_i = \frac{W_i^{increase}}{\left\|W_i^{increase}\right\|}$
Step 5: Else

Step 6:$W_i^{increase} = W_i - \sum_{j=1}^{i=1} W_i^S W_a W_a$
Step 7:$W_i = \frac{W_i^{increase}}{\left\|W_i^{increase}\right\|}$
Step 8: End if;

Step 9: If not joined, again continue the process from Step 2.

Step 8: Otherwise continue from step 1 with $increase = i + 1$
While waiting for extracting all components.

The main objective of FICA algorithm is to remove the repeated features. First of all, initialization of the weight valueW_i of all the network parameters is done randomly from 1, 2…to the current channel value, which is required for analyzing the channel performance.

The update channel weight value is calculated based on the mathematical function and it calculates the current channel weight to proportion value of update weight. The updated weight value is calculated by current channel weight with the adjacent channel weights.

This process is continued till the reach of convergence value and it is repeated to attain a maximum weight based on channel intensity. Therefore, the decision making is taken as maximum weight of feature.

B. Feature Extraction

The extraction of features is finding a set of attributes that represents an observation. The feature extraction is classified into two categories namely unsupervised and supervised FE which is used to reduce the dimensionality in order to enhance the classification accuracy.

FE plays a major role in removing the unwanted attributes which are present in pre-processed output.

C. PCCS Optimization

Optimization technique plays on the major role to select the best feature effectively. The Particle Connected Cuckoo Search algorithm is used to select the optimized features from training feature set of dictionary networks. The cuckoo bird is the major role to develop the cuckoo search algorithm, and it represents the meta-heuristic algorithm. The cuckoo bird lays their eggs in other host bird's nest. In case the host bird identifies the egg in which it can abandon the nest and builds a new nest or merely throw the egg away. Hence, the optimal result is obtained from this algorithm known as a new solution. The main aim of the algorithm is to find the priority depends on user data rate and the idle user or no transmission user eliminated based on this algorithm. Based on this, the best optimized features are selected. In algorithm II the novelty is presented to find the fitness function based on user status. The priority user is identified based only on calculation by counting the same range of user's data rate application. Therefore the idle users are eliminated to reduce the transmission delay and quality of service. The feature data matrix is passed to the optimization algorithm. First of all, finding the enable users in the feature matrix that are set to be idle or enable status. The objective function *f(x)* is used to generate the initial population of host nest such as attribute status, user status, and the utilized bandwidth information. Derive the fitness value of each channel and then sort the user's based status to obtain the perfect rank of cuckoo eggs. The status value reveals that 0 represents non-active users and 1, 2, and 3 represent active users. Then randomly generate the new solution by Levy's flight coefficients. After that, predict the next nest move and the corresponding location in step by step process. Hence, the new solution is derived and calculates the corresponding fitness value. If new fitness value is less than the old fitness value, then it is needed to replace the new fitness solution as j or else end the function. Then, removing the disabled users based on the probability function and then updates the new ones to estimate the corresponding fitness value for ranking the new solution. Therefore, the obtained features are the best features set of the algorithm. It is the optimized result of the PCCS algorithm. The algorithm to perform the optimization technique is listed as follows:

Algorithm II: Particle Connected Cuckoo Search (PCCS)

Input: Feature data matrix $\text{Feature}_{\text{data}}$
Output: Best Feature Matrix Best_{fea}

Step1: Objective function f(x)
$\text{Attribute}_{\text{sattus}}$=Strfind ($\text{Feature}_{\text{data}}$, 'Status')
$\text{User}_{\text{status}} = \text{Attribute}_{\text{sattus}}(1{:}\,n, :)$,
// n − number of the user
$\text{UE}_{\text{id}} = \text{find}(\text{User}_{\text{status}} == \text{used})$
// used − using bandwidth
Step2: In n host nest, create the first length of the user.
N_{host} =length ($\text{User}_{\text{status}}$)
Step3: Rank the eggs,
$\text{Fitness}_{\text{minimum}} = \min\ (\text{UE}_{\text{id}})$
$\text{Cuckoo}_{\text{rank}} = \text{sort}(\text{User}_{\text{status}})$
Step4: $\text{Convergence}_{\text{value}} = 10^{-5}$
// $\text{Convergence}_{\text{value}}$ − Stop Criterion
While $\text{Fitness}_{\text{minimum}} > \mathit{Convergence}_{\text{value}}$
Step5: Generate cuckoo new solution randomly by Levy Flights,
$\beta = 1.5$, $\text{Solution}_{\text{new}} = \text{User}_{\text{status}}$
$\sigma = \text{gamma\_function}(1+\beta) * \sin(\pi * (\beta/2))/\text{gamma\_function}((1+\beta)/2) * ((\beta-1)/2)^{(1/\beta)}$
$\text{Cuckoo}_{\text{x-coor}} = \text{randn}(\text{size}(\text{nest}, 2)) * \sigma$
$\text{Cuckoo}_{\text{y-coor}} = \text{randn}(\text{size}(\text{nest}, 2))$
$\text{Initial}_{\text{cuckoo-location}} = \text{Cuckoo}_{\text{x-coor}} / \text{Cuckoo}_{\text{y-coor}}^{(1/\beta)}$

$\text{Movement}_{\text{steps}}$= 0.01* $\text{Initial}_{\text{cuckoo-location}} * ()$
$\text{Solution}_{\text{new}} = \text{Solution}_{\text{new}} + (\text{Movement}_{\text{steps}} * \text{randn}(\text{size}(\text{Solution}_{\text{new}})))$
Step6: Evaluate fitness with new solution,
$\text{Fitness}_{\text{val}_{\text{new}}}$ = sum $((\text{Solution}_{\text{new}} - 1)^2)$
Step7: if $\text{Fitness}_{\text{val}_{\text{new}}} < \mathit{Fitness}_{\text{val}_{\text{prev}}}$
Replace j by the new solution
$\text{Fitness}_{\text{val}_{\text{prev}}} = \text{Fitness}_{\text{val}_{\text{new}}}$
End if.
Step 8: Empty the nest with probability Pa and update new nest.
Pa=0.25;
K=rand(size(nest))>Pa;
$\text{Steps}_{\text{forward}} = \text{rand} * \text{nest}(\text{randperm}_{\text{idx}}) - \text{nest}(\text{randperm}_{\text{idx}})$
$\text{Update}_{\text{nest}} = \text{nest} + (\text{Steps}_{\text{forward}} * \text{K})$
Step9: Evaluate Fitness with $\text{Update}_{\text{nest}}$ and rank the eggs,

End while
Step10: Find current best,

D. MLNN Classification

The optimized features are the required source of the classification method. The multi level learning is widely used in the neural networks. The set of giving exemplary decisions given to define the decision boundaries in input space is the MLNN and it's called as supervised NN(Neural Network).NN is based on three layers, which is classified as input layer, output layer, and hidden layer. In this neural network the novelty is presented and the back propagation learning algorithm is used to avoid the error between output and desired response of network.

Here, distance between user location and located tower are calculated to perform the handoff process. Then the kernel parameter based fitness function is evaluated to update the weight. The kernel trick to transform the data finds an optimal boundary between the possible outputs and it provides the accurate classification results. The equation shows that updating of weights [13] in each neuron is as following,

$$W_{let}(t+1) = W_u(t) + a(n)\, p(n)[Y^s - W_u(t)] \quad (1)$$

Here,

$$0 < a(n) < 1, p(n) = \begin{cases} d_c(n) & if\ correct\ classification \\ -1 & if\ wrong\ classification \end{cases} \quad (2).$$

Therefore,

$$d_c(n) = \frac{E_c(n)}{P_c(n)} \quad (3)$$

Where,

$E_c(n)$ and $P_c(n)$ represents the number of correct and wrongly classified patterns respectively. This equation is utilized in the proposed work to predicting the new solution weight in step 4.

TheB_f, $Target\ Label\ target_{label}$and number of neurons are the required input of MLNN classification method. Initialize the number of neurons, iteration count, and resolution count. Then, generate the target based true class, false class value and update the best feature weights of the neuron. After that, train the neural network based on the updated weight, individual neuron and the best features set. Then, Eigen matrix is calculated by corresponding Eigen value and Eigen vector. This feature is again re-updated in the kernel pattern and weight values are further processed in the classification stage. Next, derive the updated weight based on the re-update individual neuron which is used to classify the class value. Finally, the classifier list such as $Class_{out}$is displayed based on the output class. It is connected to the neighbouring cells based on estimated minimum distance. The algorithm to perform the classification technique is listed as follows:

Algorithm III: Multilevel Neural Network (MLNN) classification

Input:Best_{fea}, Target Label $\text{target}_{\text{label}}$, Number of Neuron
Output: Output class, $\text{Class}_{\text{out}}$

Step 1: Initialization steps,
$I_n =$ Number of Neuron, //$I_n -$ $\text{Individual}_{\text{neuron}}$
$\text{Maximum}_{\text{iteration}} = 10, \text{Resolution}_{\text{maximum}} = 12$
Step 2: Generate true class and false class based on the target,
For i=1:5
Else
End
End for
Step 3: Weights of the neuron is updated,

$$\text{Negative}_{\text{points}} = \max(\text{falseclass}),$$

$$\text{Max}_{\text{val}} = \max(B_f),\ \text{//Where, } B_f - \text{Best}_{\text{fea}}$$

$$\text{Min}_{\text{val}} = \min(B_f).$$

Interval= $\text{Max}_{\text{val}} - \text{Min}_{\text{val}}$

$$U_w = 10 * \frac{(\text{rand}(\text{size}(B_f,2), I_n) - 0.5)}{\text{Interval}/\text{size}(B_f,2)} \text{//Where, } U_w - \text{Update}_{\text{weight}}$$

Step 4: Train the neural network with $\text{Update}_{\text{weight}}$, I_n and B_f

Initialize, $\text{eigen}_{\text{vec}} = 0$ $\text{Kernel}_{\text{param}} = 1$

For n=1 to , I_n
$\text{W}_{\text{let}} = U_w(:,:,\text{n})$
Calculate eigen matrix,
$\text{eigen}_{\text{val}} = \text{eig}(\text{W}_{\text{let}} * \text{W}_{\text{let}}')/{I_n}^2$

$$\text{eigen}_{\text{vec}} = \frac{[\max(\text{eigen}_{\text{val}}) + \min(\text{eigen}_{\text{val}})] * \text{eigen}_{\text{vec}}}{\text{size}(B_f, 2)}$$

Fitness (n) = $\text{eigen}_{\text{vec}}$

End for
Step 5: Update
$\text{Kernel}_{\text{param}}$ and reupdate weight$\text{Reupdate}_{\text{weight}}$
$\text{Kernel}_{\text{param}} = \text{Kernel}_{\text{param}} + 1$
$\text{mean}_{\text{fit}} = \text{mean}(\text{fitness})$
$[\text{min}_{\text{fit}}, \text{Kernel}_{\text{param}}] = \min(\text{fitness})$

$$\text{Reupdate}_{\text{weight}} = U_w(:,:,\text{Kernel}_{\text{param}})$$

Step 6: Again $U_w$based on the individual neuron,
For k =1:I_n
r=rand
k1 = ((round $(I_n - 1)*(e^{R_n * r}) - 1)/(e^{R_n} - 1)) + 1$
r=rand
k2 = ((round $(I_n - 1)*(e^{R_n * r}) - 1)/(e^{R_n} - 1)) + 1$
r=rand (I_n, size (B_f, 2))

$$U_w = 1 + 0.6 * (\text{rand} - 0.5) * \text{Reupdate}_{\text{weight}}(:,:,\text{k1}) * \text{r} + \text{Reupdate}_{\text{weight}}(:,:,\text{k2}) * (1 - \text{r})$$

End for
Step 7: Return class with U_w
If ~ is emptyU_w
End if

IV. Experimental Analysis

In mining process, many processes provide required classification and retrieval of data by performing matching up of data from library. These are solely based on matching trained dictionary with that of the predicted data; the classification of user data for channel scheduling is based on the training features of the dictionary. As discussed earlier a novel feature extraction method and classification for data mining process is done to retrieve information about a network for LTE scheduling. The data mining technique for this proposed work can be done by stages the first stage is preprocessing stage, which provides dimensionality reduction in the given input feature data during the training case. The second stage is feature extraction algorithm which extracts the features of the network and its degree of angle with the distance between the features set. In this stage, we have implemented and worked through Fast Independent component analysis (FICA) method which reduces the redundant features and extracts the relevant network features and its degree of the angle with distance based estimation. Table II shows the user attributes taken for analysis and the distance calculation of each user from different base station is done. Both the feature identification and structure of network framing increases the performance of data mining analysis. Fig 2 shows the reduced user and cell deployment.

Table III tabulates the updated features after estimating distance after dimensionality reduction. The third stage is an optimization technique for selecting the best features in the training feature set. Hence for the optimized approach of selection and classification of user attributes and labelling, we have implemented using particle cuckoo search (PCCS) for selection and then according to that extracted features, we perform classification using multilevel neural network (MLNN) to predict the category to which each individual customers belong. Also for providing best customized channel for the users by the speed of data rate used by them neural network learning is needed.

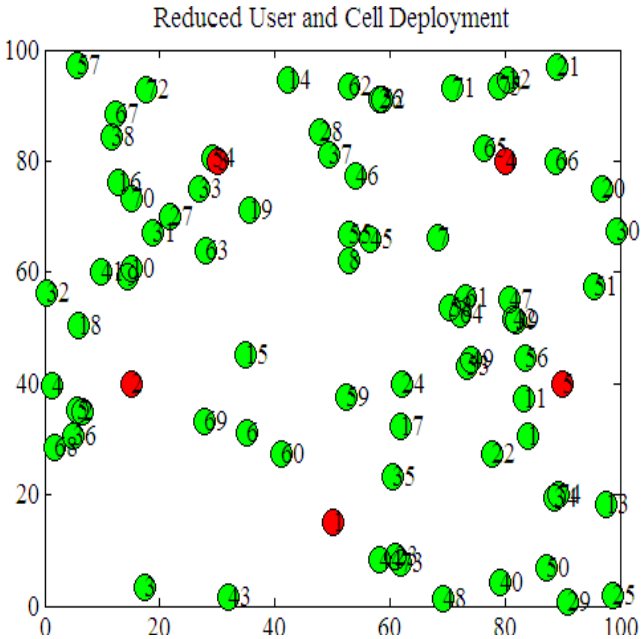


Fig. 2. Reduced User and Cell Deployment

Here the multi-level classification is initially provided to categorize. After categorizing, it also predicts the label of network category from the given network information. Then this information is passed to LTE scheduling system for best communication. Fig 3 shows the final channel optimized for users. This type of feature extraction and classification will provide better performance rate comparing with traditional methods. Thus the performance of the system is increased by using the data mining techniques. These steps are not mandatory all the time for scheduling of channels in the LTE network. However if these steps are performed we will get a better performance system and high speed communication without delay between the users.

V. PERFORMANCE ANALYSIS

This section presents the performance analysis and evaluation of the proposed system based on some measures such as true positive, true negative, false positive, false negative, specificity, accuracy, recall, precision, F-measure, and sensitivity. These measures are considered to be the generalized analysis of the proposed model. Table IV shows the data and description of the attributes [12] which are utilized in the proposed work and their descriptions. Here, the 100 customer ID is connected to 4 different cells. It takes 1-5GHz carrier frequency and 1-3ms for their updating time. The data usages are measured as bits per second and bandwidth in terms of MHz. the user utilized the equipment speed as 3, 30,120km\h.

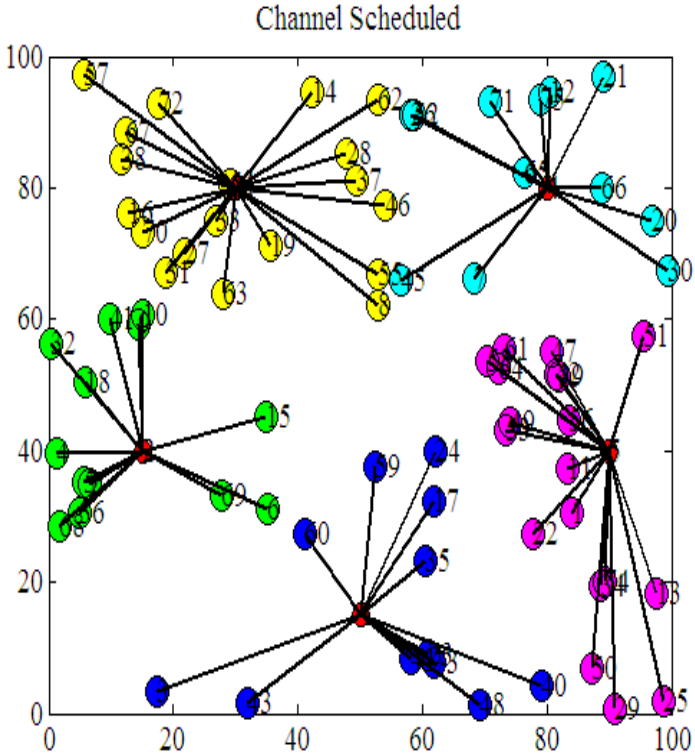


Fig. 3. Scheduled Channels

Table II. User attributes taken for analysis

Customer ID	Carrier Freq (GHz)	Update Time of CD(ms)	Hand Over Delay(ms)	UE Speed (km/h)	eNB Power (dBm)	Distance 1	Distance 2	Distance 3
1	5	1	3	99	32	37.508	45.821	21.358
2	3	3	26	25	32	46.890	56.238	42.123
3	5	2	21	17	32	20.450	52.369	54.639
4	4	3	22	99	32	28.752	44.236	20.215
5	3	1	20	78	32	50.263	49.562	16.582
6	4	2	16	4	32	19.235	31.542	20.654
7	5	3	10	104	32	58.236	26.542	52.321
8	1	3	20	63	32	50.562	42.265	47.589
9	2	2	3	67	32	33.333	17.236	58.456
10	5	1	15	52	32	48.594	36.528	44.213

Table III. Updated features after estimating distance

1	2	3	4	5	6	7	8	9	10	11	12	13	14
Customer ID	Carrier Freq (GHz)	Update Time of CD(ms)	Hand Over Delay (ms)	UE Speed (km/h)	eNB Power (dBm)	Bandwidth (MHz)	Hand off status	Data usage (Mbps)	Distance 1	Distance 2	Distance 3	Distance 4	Distance 5
1	5	1	3	99	32	4	3	1094	31.589	70.5985	81.985	63.7968	25.2645
2	3	3	26	25	32	5	3	1386	76.1426	85.2456	60.5819	10.5986	39.4927
3	5	2	21	17	32	5	1	3	40.631	9.1719	51.825	83.1780	77.8095
4	4	3	22	99	32	3	2	308	56.036	77.1562	9.1719	60.1578	65.870
5	3	1	20	78	32	3	3	1213	13.251	53.625	71.652	66.675	26.659
6	4	2	16	4	32	2	0	0	60.591	18.723	66.775	87.075	59.6325
7	5	3	10	104	32	4	1	20	24.8647	85.7737	96.1255	81.439	7.3217
8	1	3	20	63	32	4	0	0	20.256	5.112	21.7277	26.3154	38.127
9	2	2	3	67	32	4	1	10	70.152	97.426	77.9283	78.1254	35.2403
10	5	1	15	52	32	2	2	315	76.3477	82.5066	87.3670	13.623	47.2335

Analysis and Measures

The term sensitivity measure is the fraction of true positive to the total number of false negative and true positive. It is calculated as follows.

$$Sensitivity = \frac{TP}{count\ of\ TP + count\ of\ FN} \quad (4)$$

The word specificity is well defined as the division of true negative to the summation of false negative and true negative measure. It is calculated as follows,

$$Specificity = \frac{TN}{number\ of\ TN + number\ of\ FP} \quad (5)$$

The F-measure is calculated based on the values of precision and recall. The mathematical formulation of F-measure is calculated as

$$F - measure = \frac{precision.recall}{precision + recall} \quad (6)$$

The term accuracy is the proportion of the total of true negative and true positive to the summation of all TP, TN, FP and FN. It is calculated as follows.

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \quad (7)$$

Precision is the measure of the proportion of true positive to the sum of false positive and true positive. It is called as true positive rate and its represent as,

$$Precision = \frac{T_p}{T_p + F_p} \quad (8)$$

Recall is defined as the proportion of true positive to the sum of true positive and false negative rate. It is considered as follows,

$$Recall = \frac{T_p}{T_p + F_N} \quad (9)$$

Table V presents the performance analysis of the proposed system. The evaluation of the proposed system performance based on some measures such as true positive, true negative,

Table IV . Data Description

Attributes	Description
Customer Id	Unique customer ID's connected to different cells.
Carrier Frequency	Carrier frequency value in Ghz (1 -5 Ghz).
Update time of CD	Updating time of customers' data in ms (1 -3 ms).
Handover delay	Delay time in handoff procedure in ms(0-30 ms)
Handoff Status	Whether a successful handoff or not (Y/N or 0/1).
UE speed	User equipment speed in km/h (3, 30,120 km/h).
eNodeB	Element of an LTE radio access network power in decibel meter.
Bandwidth	Bandwidth in terms of MHz
Data Usage	Data usage in terms of bits per second

false positive, false negative, specificity, accuracy, recall, precision, F-measure, and sensitivity.

The result provides 100% of sensitivity, 97.43% of F-measure, and 98.21% of specificity. From the results, it is observed that the proposed classification technique provides the best results. The result provides 98.67% of accuracy, 95% of precision and 100% of recall. From the results, it is detected that the proposed classification technique yields better results than existing methods effectively.

Table V. Performance Measures

Parameters	MLNN
True Positive	19
True Negative	55
False Positive	1
False Negative	0
Accuracy (%)	98.67
Sensitivity (%)	100
Specificity (%)	98.21
Precision (%)	95
Recall (%)	100
F Measure (%)	97.43

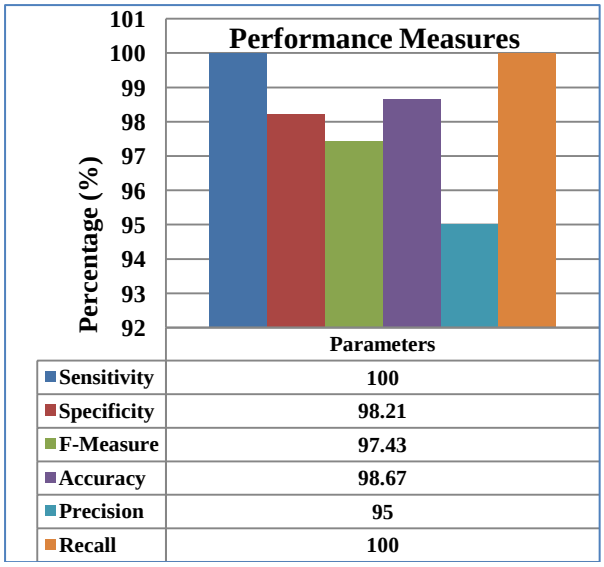


Fig. 4. Performance Measures

VI. Comparative Analysis

This section deliberates the efficacy of the proposed system by comparing with the existing classification and optimization methodologies such as naïve Bayes, stacking C, DTNB, random forest, J48, ridor, decision table, zero R, grading and simple cart techniques in terms of correctly classified instances, in corrected classified instances, kappa statistics, root mean squared error (RMSE), time to build.

A. Correctly and Incorrectly Classified Instances (ICI)

This error metric is to evaluate the performance of incorrectly classified instances of proposed with the existing methods.(Fig 5 &6) The low value of the error ratio indicates the efficiency of proposed work. From the comparison, the random forest method provides less ICI (%) associated with other existing systems which is 2.02%. The comparative analysis between the proposed and the existing random forest system depicts that the proposed MLNN system provides 1.34%. Hence, it is clearly shown that the proposed MLNN yields a 33.66% reduction in the error percentage respectively.

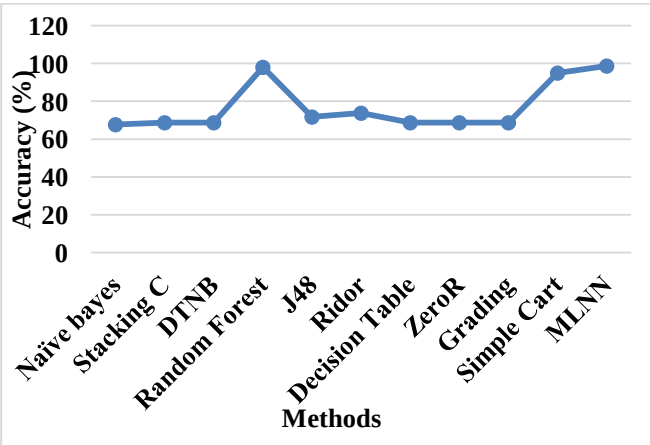
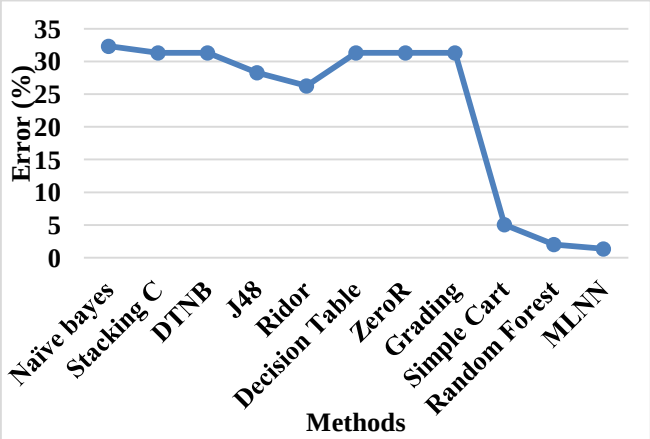


Fig. 5 Correctly Classified Instances

Fig. 6 Incorrectly Classified Instances



B. Kappa Statistics

Kappa measures the percentage of data meets the user requirement. It is always less than or equal to one. If it is, 1 represent a perfect agreement or < 1 represent a small amount of perfect agreement. From the comparison, the random forest method provides a high kappa coefficient related to other existing systems which is 0.9522. The comparative analysis between the proposed and the existing random forest system depicts that the suggested MLNN system provides 0.9966 values. Hence, it is clearly shown that the proposed MLNN yields almost identical to one for perfect agreement.

C. Root Mean Squared Error (RMSE)

The RMSE value is the measure of the difference between the measured value and the observed value. It lies in between 0 and 1. From the comparative analysis, the simple cart method provides the less RMSE value associated with other existing systems which is 0.2052. The comparative analysis between the proposed and the traditional simple cart system depicts the proposed MLNN system provides 0.1333. Hence, it is clearly illustrated that the proposed MLNN system RMSE value yields a 35% reduction than the existing system respectively.

Fig. 7 Kappa Statistics Analysis

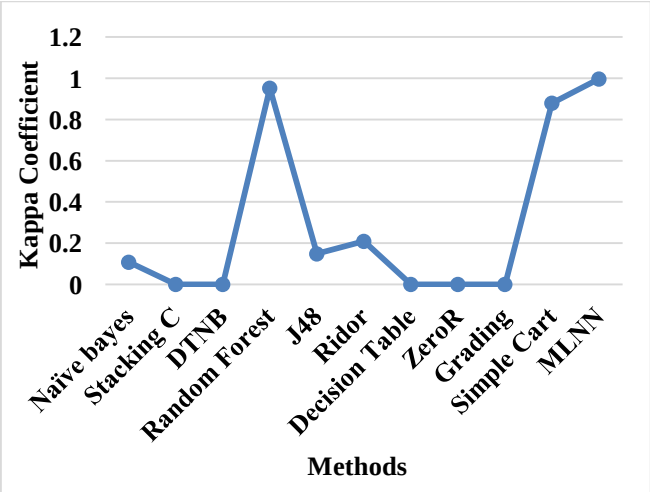
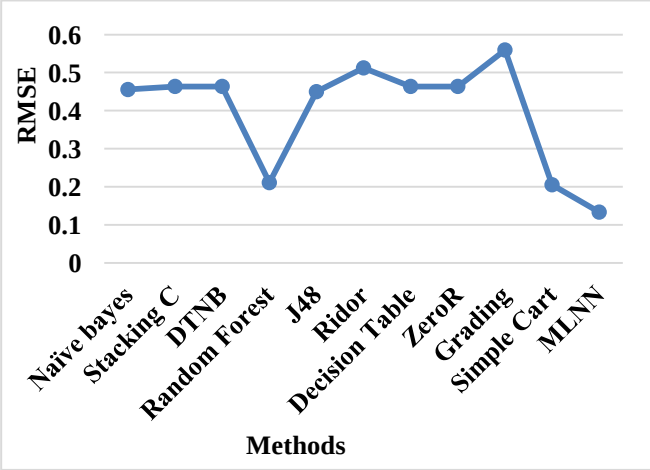


Fig. 8 Root Mean Squared Error Rate



D. Build Time

Build time is the measured time to transfer the information to other users. The proposed MLNN system compared with the tree based structure system; proposed work provides greater performance than existing. The proposed system simultaneously performs the optimal channel capacity with the time-based transmission. From the comparison, the proposed system provides 0.04 second. The build time analysis of the proposed system provides better performance than the existing work. This work considers both of channel capacity and time.

Fig. 9 Build Time Analysis

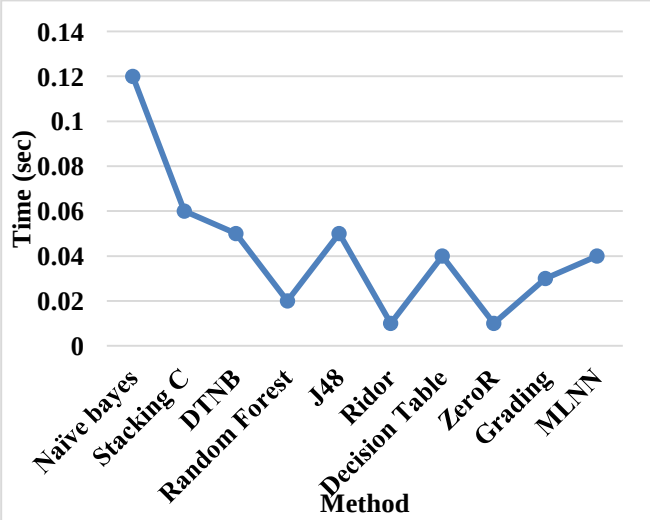


Table VI shows the summary of comparison of different [12] [16] with the parameters are correctly classified instance, incorrectly classified instance, kappa statistic, root mean squared error and build time.

Table VI. QOS Based Comparison of Existing and Proposed Method

Algorithms	Correctly Classified instance (%)	Incorrectly Classified instance (%)	Kappa statistic	Root mean squared error	Time to Build
Naïve Bayes	67.6768	32.3232	0.1076	0.4556	0.12

stacking	**68.6869**	**31.3131**	**0**	**0.4638**	**0.06**
DTNB	**68.6869**	**31.3131**	**0**	**0.4638**	**0.05**
Random Forest	**97.9798**	**2.0202**	**0.9522**	**0.2108**	**0.02**
J48	**71.7172**	**28.2828**	**0.1481**	**0.45**	**0.05**
Ridor	**73.7374**	**26.266**	**0.209**	**0.5125**	**0.01**
Decision Table	**68.6869**	**31.3131**	**0**	**0.4638**	**0.04**
ZeroR	**68.6869**	**31.3131**	**0**	**0.4638**	**0.01**
Grading	**68.6866**	**31.31**	**0**	**0.5569**	**0.03**
Simple cart	**94.9491**	**5.0505**	**0.8794**	**0.2052**	**0.03**
MLNN	**98.66**	**1.34**	**0.9968**	**0.1333**	**0.04**

VII. CONCLUSION

Forecasting the relevant channel and channel scheduling services based on the channel information and user availability in the network is major task in recommendation model. This paper discussed the limitations of channel scheduling such as data complexity, channel capacity and available in recommendation models. To overwhelm this limitation, proposed the feature extraction and classification method for data mining process. The Fast Independent Component Analysis (FICA) reduces the dimensionality highly. The feature extraction utilized the handoff attributes to identify the relevant feature and remove the redundant features based on the distance calculation. Effectively, then select the best feature dataset from the dictionary networks by using the Particle Connected Cuckoo Search (PCCS) optimization algorithm. According to that extracted features, perform Multilevel Neural Network (MLNN) classification method to classify the channel category. Finally, the information is passed to LTE scheduling system for better communication. The comparative analysis between the proposed MLNN with the existing naïve Bayes, stacking C, DTNB, random forest, J48, ridor, decision table, zero R, grading and simple cart in relationship with correctly classified instances, in correcting classified instances, kappa statistics, root mean squared error (RMSE), and build time analysis confident that the efficacy of the MLNN classification system.

REFERENCES

1. Ang EM, Wee K, Pang YH and Phang KK , "A performance analysis on packet scheduling schemes based on an exponential rule for real-time traffic in LTE," EURASIP Journal on Wireless Communications and Networking,vol 1,2015, pp.2-12.
2. Baghi S and Farzanegan MD , "A novel delay based scheduling algorithm for video traffic in LTE," 2nd International Conference Knowledge-Based Engineering and Innovation (KBEI),2015,pp. 514-520.
3. Bi M, Xiao S, Li J and He H , "A bandwidth-efficient channel allocation scheme for mitigating FWM in ultra-dense WDM-PON," Optik-International Journal for Light and Electron Optics, vol. 125,no. 8, 2014,pp. 1957-1961.
4. Bukhari. S. H. R , Rehmani. M. H, and Siraj. S , "A survey of channel bonding for wireless networks and guidelines of channel bonding for futuristic cognitive radio sensor networks",*IEEE Communications Surveys & Tutorials*,(18) ,2016 , pp.924-948.
5. Chen T, Zhao X, Gao T and Zhang L, "A novel downlink scheduling strategy for traffic communication system based on TD-LTE technology,"SpringerPlus, vol. 5,no.1,2016, pp. 1631.
6. Ding Z, Liu Y, Choi J, Sun Q, Elkashlan M, Chih-Lin I, "Application of non-orthogonal multiple access in LTE and 5G networks", IEEE Communications Magazine, vol. 55,no. 2,2017, pp. 185-191.
7. Erpek.T ,Abdelhadi.A, & Clancy.T.C, "Application-aware resource block and power allocation for LTE", *Systems Conference (SysCon) Annual IEEE* ,2016,pp.1-5.
8. Feng M, Mao S and Jiang T, "Joint duplex mode selection, channel allocation, and power control for full-duplex cognitive femtocell networks", Digital Communications and Networks, vol. 1,no. 1, 2015,pp. 30-44..
9. Ferdosian N, Othman M, Ali BM and Lun KY, "Greedy–knapsack algorithm for optimal downlink resource allocation in LTE networks,"Wireless Networks, vol. 22,no. 5,2016, pp. 1427-1440.
10. Gao C, Sheng X, Tang J, Zhang W, Zou S and Guizani M, "Joint mode selection, channel allocation and power assignment for green device-to-device communications," 2014 IEEE International Conference on Communications (ICC), 2014,pp. 178-183.
11. Grondalen O, Zanella A, Mahmood K, Carpin M, Rasool J and osterbo , "Scheduling policies in time and frequency domains for LTE downlink channel: a performance comparison," IEEE Transactions on Vehicular Technology, vol. 66,no.4,2017, pp.3345-3360.
12. Hans. A, Singh. N, & Kalra. S , "Advanced Data Mining Appraoch For Handoff Procedure's in Lte Technology,""International Journal of Database Theory and Application,9(2), pp.9-16.
13. Kashtiban AM, Razmi H & Kozehkonan MK , "Combined LVQ neural network and multivariate statistical method employing wavelet coefficient for EEG signal classification," 2011 IEEE International Conference on Mechatronics (ICM), 2011,pp. 809-814.
14. Lei L, Yuan D, Ho CK and Sun S , "A unified graph labeling algorithm for consecutive-block channel allocation in SC-FDMA," IEEE Transactions on Wireless Communications, vol. 12,no. 11,2013, pp. 5767-5779.
15. Li M, Zhao L and Liang H, "An SMDP-based Prioritized Channel Allocation Scheme in Cognitive Enabled Vehicular Ad Hoc Networks," IEEE Transactions on Vehicular Technology,vol.66 (9),2017,pp.7925-7933.
16. Mohan.D & Amalanathan, G. M., "A Survey on Long Term Evolution Scheduling in DataMining," *Wireless Personal Communications* Vol.*102*. *No*.3,2018,pp.2363-2387.
17. Nardini G, Stea G, Virdis A, Sabella D and Caretti M, "Practical large-scale coordinated scheduling in LTE-Advanced networks", Wireless Networks, vol. 22,no. 1,2016, pp. 11-31.
18. Nsiri B, Nasreddline M, Ammar M, Hakimi W and Sofien M , "Modeling and performance evaluation of scheduling algorithms for downlink LTE cellular network," The Tenth International Conference on Wireless and Mobile Communications,2014,pp.60-64.
19. Song. L, Niyato. D, Han. Z & Hossain. E , "Game-theoretic resource allocation methods for device-to-device communication," *IEEE Wireless Communications,(21),2014,pp.*136-144.
20. Trivedi. R. D and Patel. M , "Comparison of Different Scheduling Algorithm for LTE , " *International Journal of Emerging Technology and Advanced Engineering, 4*(5),2014, pp.334-339.
21. Tung L-P, Lin Y-D, Kuo Y-H, Lai Y-C and Sivalingam KM, " Reducing power consumption in LTE data scheduling with the constraints of channel condition and QoS," Computer Networks, vol. 75, 2014, pp. 149-159.
22. Wang Y-C and Tsai T-Y, "A Pricing-Aware Resource Scheduling Framework for LTE Networks," IEEE/ACM Transactions on Networking,vol.25 (3), 2016 , pp.1445-1458.

AUTHORS PROFILE

Divya Mohan pursuing her research in the Department of CSE at VIT University, India. She received M.E degree in communication systems from Velammal college of Engineering and Technology, India. She has published more than three papers in reputed journals and conferences. Her current research interests include Wireless networks, Machine Learning, LTE Scheduling and soft computing.

Geetha Mary A received her Ph.D. from VIT University, Vellore, India. She has completed M.Tech in Computer Science and Engineering from VIT University and B.E. from University of Madras, Tamil Nadu, India. She is working for VIT University as Associate Professor. She was awarded Merit Scholarship for her best academic performance for the year 2007-2008 during her M.Tech study. She has authored more than 20 journal and conference papers. She has authored book chapters in the area of data mining and artificial intelligence. Her research interests include security for data mining, databases and intelligent systems. She works to empower healthcare management using computer science. Dr. Geetha Mary is associated with many professional bodies like IACSIT, CSTA and IAENG.