An Effective Demand based Optimal Route Generation in Transport System using DFCM and ABSO Approaches

Archana M. Nayak[1]*
Research Scholar
Computer Engineering Department
Gujarat Technological University, Gujarat, India

Nirbhay Chaubey[2]
Associate Professor
Dean of Computer Applications
Ganpat University, Gujarat, India

***Abstract*—The transportation network service quality is generally depends on providing demand based routing. Different existing approaches are focused to enhance the service quality of the transportation but them fails to satisfy the demand. This work presents an effective demand based objectives for optimal route generation in public transport system. The importance of this work is providing demand based optimal routing for large city transportation. The proposed demand based optimal route generation process is described in subsequent stages. Initially the passengers in each route are clustered using Distance based adaptive Fuzzy C-means clustering approach (DFCM) for collecting the passengers count in each stop. Here the number of cluster members in each cluster is equivalent to the passenger count of each stop. After the clustering process, adaptive objectives based beetle swarm optimization (ABSO) approach based routing is performed with the clustered data. Then re-routing is performed based on the demand based objectives such as passenger's count, comfort level of passengers, route distance and average travel time using ABSO approach. This ABSO approach provides the optimal routing based on these demand based objectives. The presented methodology is implemented in the MATLAB working platform. The dataset used for the analysis is Surat city transport historical data. The experimental results of the presented work is examined with the different existing approaches in terms of root mean square error (9.5%), mean error (0.254%), mean absolute error (0.3007%), correlation coefficient (0.8993), vehicle occupancy (85%) and accuracy (99.57%).**

Keywords*—Clustering; optimization; demand based objectives; comfort level; optimal routing

I. Introduction

The increased population rate in large cities is one of the major transportation issues in developing countries. The unplanned organization in those cities increases the mobility demand and improves the tendency of private vehicle transport ownerships [1, 2]. This issue is common for all large cities have to deal with hundreds of vehicles running through their street each day. The congestion of traffic is increasing day by day even in rural areas also. Moreover, the traffic congestion associated delays are happening more often [3, 4]. A public transport network plays an important role in changing the demand from private to effective transportation and reduced traffic congestion [5]. The better public transportation can avoid the traffic congestion. To enhance the cooperation in public transport systems, different demand based models are introduced. Passenger comfort [6] is one of the vital indexes used to validate the quality of public transport system.

The quality of the public transport system is depends on increasing the passenger comfort. Therefore, enhancing the bus comfort is a great attention by the public transport to attract the passengers [7-9]. The passenger's choice of transport is depends on various factors like comfort ability, travelling time, expense and reliability. Therefore, transport authorities are making attention to attract the passengers by providing passengers demand based transportation [10, 11]. The comfort of transportation is categorized into two classes. One is based on the performance of vehicle and the other one is depends on the operation of transport. Various studies are presented for providing demand based public transportation [11]. The investigated works states that uncertainties in travel times and traffic increases the passenger's stress level and reduces the passenger's quality [12, 13]. Some of the researchers proved that the comfort is depends on the passenger's load. So, accurately determining the bus comfort and providing demand based effective bus routing is important for improving the quality of transportation [14]. The load factor of passenger's and the bus travel time is an important and foremost for the quality of transportation [15]. The proficient public transport system can lessen the negative issues of private transport network.

Moreover, the quality of transport is mainly depends on minimization of travel time and waiting time. This effectively increases the performance of transportation [16]. The transportation authorities are responsible for generating the optimal routes. Based on the different objectives, transport authorities are determining the demanded bus routes and schedules [17]. The main concern of public transport is satisfying the public needs and the transport authorities are focussing on their profit. Moreover, different constraints are considered in the existing researches for the generation of transportation routes [18]. But the optimal route generation is still in demand.

Numerous optimization approaches are utilized in the existing works to enhance the transportation framework. Some of the existing optimization approaches are, genetic algorithm

*Corresponding Author.

(GA) [19, 20], simulated annealing (SA), Tabu search (TS) and Swarm intelligence (SI) [21]. These optimization approaches are proved their capability, effectiveness in optimal route generation in transportation network. But all the major issues of transportation are not solved as a whole and optimal route generation is still in demand. Moreover, demand based optimal route generation is not performed in the existing works.

The main motivation of proposed work is providing demand based transportation in large cities. Therefore, the quality of transportation can be achieved by considering major objectives such as passenger's count, comfort level of passengers, route distance and average travel time. In the presented work, demand based optimal route generation is proposed to improve the quality of transportation network. The major contributions of the presented methodology is described as,

- To effectively cluster the passengers in the transportation network, distance based fuzzy c-means clustering (DFCM) approach is utilized. This approach is utilized to gather the accurate passenger's count of each station in the transportation network.
- Optimal routing is generated by using effective objective based adaptive beetle swarm optimization. Re-routing is performed based on the different objectives such as comfort level of passengers, passenger's count, route distance and average travel time.
- Demand based effective routing is obtained by the proposed ABSO optimal routing approach. Here, the performance of the proposed approach is enhanced based on the considered effective objectives. This enhances the quality of the transportation by providing demand based vehicles.

The organization of this research work is described as: Section 2 surveys the recent related works, the presented approach is described in section 3, results and their discussion is provided in section 4 and the section 5concludes the paper respectively.

II. Related Work

Ovidiu Cosma et al. [22] presented an innovative methodology for two-stage transportation issues with fixed costs related to the routes. The developed heuristic algorithm solves the investigated transportation issues and it was attained by integrating the linear programming optimization issue within the genetic optimization process. Moreover, an effective local search process was incorporated with genetic optimization algorithm able to enhancing the global search. The proposed methodology effectively solves the transportation issue and enhances the quality in transportation. The performance can be enhanced further by using other optimization approaches in transportation.

Jiangtao Liu et al. [23] developed a routing based on the integration of travel demand, vehicle supply and restricted infrastructure. Based on these constraints, vehicles were optimally assigned in the routes to satisfy the passenger's. Various decomposition techniques were adapted in this research. Vehicle based integer linear programming model was presented for space time state (STS) with Dantzig Wolfe decomposition methodology. The presented methodology effectively increases the efficiency and reduces the congestion in transportation. But the proposed methodology was challenging to predict the optimal routes based on the travel demand.

Jishnu Narayan et al. [24] developed an integrated bus route selection and scheduling framework for static and dynamic public transport networks. The introduced framework based on the dynamic demand based route generation using an iterative learning framework. The effective routing based on this proposed approach effectively reduces the average waiting time of passengers. The combination of fixed and flexible transport networks enhances the routing ability and satisfies the passengers. Moreover, the day to day learning was considered to effective selection of routes and allocation of transport vehicles. The proposed approach can be improved in future by incorporating the time and reliability of passenger's.

Mahmoud Owais and Abdullah Alshehri [25] developed a pareto optimal route generation in transportation networks. Instead of considering the travel time like existing approach, the presented approach considered the relation among the traffic flow and the respective trip time. Based on the demand based information, the initialized routes were changed and select the traffic allocation approach. Finally, different objectives based approach was conducted with pareto optimal route generation. Further, the research performance can be enhanced by using effective recent techniques in future.

Philipp Heyken Soares [26] developed a zone based optimized routing in an urban network. Node based optimization was introduced on transportation routes by predicting the passenger trips for each stops. Here, the hybrid methodology was utilized for evaluate the journey times of source to destination. The travel time of each stop was determined by the node based procedures. The developed optimization procedure was applied on input considered real world database and the enhancement was shown existing routing framework. Further enhancements were possible in various factors based on passenger's comfort and satisfaction.

The investigated existing approaches are tried to solve the issues in the transportation system. But still, the existing approaches are not reached the satisfaction level by solving all the issues. Some of the approaches are focused on increasing the passenger's while decreasing the comfort of passengers. Therefore, the quality of transportation is needs to be enhanced by considering major objectives. The proposed work considered the four effective objectives such as passenger's count, comfort level of passengers, route distance and average travel time to provide the demand based routing. This resolves the issues present in the existing routing. Based on the proposed demand based routing, the service quality of the transportation can be increased.

III. Proposed Methodology

This work presents an effective demand based optimal route generation in transport system. At first, Distance based adaptive Fuzzy C-means clustering approach (DFCM) is utilized for clustering the passengers in each stop. Here, the number of cluster members in each cluster is equivalent to the passenger's count of every station. Afterwards, adaptive beetle swarm optimization approach based routing is performed based on the clustered data. Further, re-routing is performed based on the effective objectives such as passenger's count, comfort level of passengers, route distance and average travel time using ABSO approach. In the optimal route generation approach, objectives are analysed in each pair of stations in the route using the ABSO approach. This ABSO approach enhances transportation system by providing demand based routing. The schematic diagram of the presented methodology is depicted in Fig. 1.

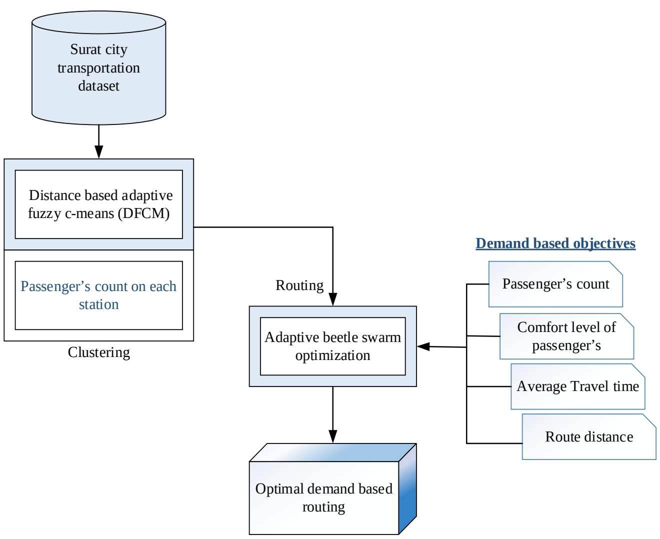

Fig. 1. Schematic Diagram of Presented Methodology.

The process of demand based optimal routing in the transportation system is described in the subsequent sections. Initially, the input Surat city transportation history data is considered and the passenger of each station is obtained by the proposed DFCM clustering. Afterwards, routing is performed by the proposed ABSO approach and demand based objectives are updated in this optimization approach to effectively performs the re-routing. This methodology effectively provides the demand based routing.

A. Distance based Adaptive Fuzzy C-Means (DFCM) Clustering

The proposed DFCM clustering approach minimizes the objective function through the iterative process. Consider, $Q = \{q_1, q_2, q_3, \ldots q_p\}$ represents the vectors of cluster centre, $\bar{U} = [\bar{\mu}_{kp}]^{m \times p_{opt}}$, $\bar{\mu}_{kp}$ represents the membership degree of input data y_k to the cluster centre q_p, $\bar{\mu}_{kp} \in (0, 1)$. The objective function of the presented DFCM approach is described as,

$$\bar{J}_{DFCM} = \sum_{p=1}^{P_{opt}} \sum_{k=1}^{m} \bar{\mu}_{kp} \left\| y_k - q_p \right\|^{N} + D(k) \tag{1}$$

Here, N represents the constant to control the fuzzy function of clustering results $N \in [1, +\infty)$. Here, the cluster membership function is updated by the subsequent condition (2),

$$\bar{\mu}_{kp} = \frac{1}{\sum_{l=1}^{P_{opt}} \left(\frac{D_{kp}}{D_{kl}} \right)^{\frac{2}{n-1}}} \tag{2}$$

The cluster centre vectors are updated by the subsequent condition (3),

$$q_p = \frac{\sum_{k=1}^{m} (\bar{\mu}_{kp})^n y_k}{\sum_{l=1}^{m} (\bar{\mu}_{kp})^n} \tag{3}$$

The fuzzy c-means clustering is processed into different steps to minimize the objective function $\bar{J}_k$ with the update of $\bar{U}$ and $\bar{Q}$.

The effectiveness of the fuzzy c-means clustering is based on the initial means. The means are having great impact on the final clustering. The existing FCM approaches are based on the arbitrary initialization to examine the centroid for clustering. But, the proposed DFCM approach centroids are selected according to the dense nodes positioning. The nearest neighbouring nodes are considered as a dense nodes. The nearest neighbouring nodes are the nodes with minimum distance. The distance of the nodes is evaluated based on the dense nodes to estimate the passenger's count in each station. Distance is evaluated by the subsequent condition (4),

$$D(k) = \sum_{l=1}^{M} \exp\left[-\left\| y_k - y_l \right\|^2 * \left(\frac{1}{Dense\ cluster} \right)^2 \right] \tag{4}$$

Here, $Dense\ cluster = \frac{M}{P_{opt}}$ (5)

Here, P_{opt} represents the optimal cluster center. Based on these evaluations, the high dense data points are considered as a first cluster centre, and the farthest from the first centre is considered as a second centre and so on. The P^{th} cluster centroid is satisfied with the condition $Maximum\left\{ \prod_{k=1}^{p-1} (q_p - q_k) \right\}$. The evaluated distance using condition (4) is updated in the condition (1). Then, the objective function satisfies the condition $\left| \bar{J}_k^{(t)} - \bar{J}_k^{(t-1)} \right| < \delta$, and then forms the optimal clustering. Here, δ represents the threshold function. The initial cluster centre is needs to be changed to overcome the issue of distance among the node position and the cluster centre is zero. The cluster centres are updated based on the consideration that if the neighbouring node centre is far from $q_1, q_2, q_3, \ldots\ldots q_p$ is then considered as the updated cluster centroid. Clusters are formed by assigning nearest sensor nodes to the cluster based on this DFCM approach. The nearest sensor nodes are considered based on the minimum distance of sensor nodes. The pseudo code of the DFCM clustering is described in algorithm 1.

In algorithm 1, the proposed DFCM clustering steps are provided. This effectively clusters the historical data of passengers. Evaluate the objective function $\bar{J}_{DFCM}$ as per the condition (1), if the condition satisfies $\left|\bar{J}_k^{(t)} - \bar{J}_k^{(t-1)}\right| < \delta$ then the optimal clustering is obtained. The passenger's count details of each station are obtained in this clustering approach.

Algorithm 1: Pseudo code of DFCM clustering

Input: Initialize the passenger data in the transportation network $S = \{s_1, s_2, \ldots\ldots s_n\}$ and the initial centres $Q = \{q_1, q_2, \ldots\ldots q_p\}$

Output: Optimal clustering

Begin

- **Initialization** of passenger's data $S = \{s_1, s_2, \ldots\ldots s_n\}$ and centroids $Q = \{q_1, q_2, \ldots\ldots q_p\}$.
- **Compute** the objective function $(\bar{J}_{DFCM})$ utilizing condition (1).
- **Evaluate** membership nodes and cluster centres with conditions (2) and (3) at $\bar{J}_{DFCM}$ computation step.
- **Compute** the adaptive distance evaluation for passenger data using the condition (4)
- **Update** the cluster centroids by utilizing the condition (3),
- **If** the clustering process attains best outcome at maximum iterations, then the iteration is terminated.
- **Else**
- **Repeat** the above procedure until reaches the best solution.
- **Return** optimal clustering

End

B. *Demand based Optimal Routing in Transportation System*

The demand based optimal routing is attained by the proposed adaptive objectives based beetle swarm optimization approach. Optimal re-routing is performed by updating the demand based objectives such as passenger's count, comfort level of passengers, route distance and average travel time are described in the subsequent sub sections.

1) Demand based objective measures: In this section, effective objectives are evaluated based on the transportation demand is described in the subsequent sections,

a) Comfort Level of Passengers: The passenger' comfort level is depends on the number of passenger's in the vehicle and the capacity of the vehicle. Comfort level of passenger is estimated by the subsequent condition (6),

$$\bar{C}_\alpha = \frac{\bar{P}_N}{\bar{V}_C} \tag{6}$$

Here, α represents the passenger's comfort level, $\bar{P}_N$ represents the number of passengers in the vehicle and $\bar{V}_C$ represents the vehicle's capacity.

b) Passenger's Count: The number of passengers in each station is computed based on the count between source and destination. Passenger's count is computed by the subsequent condition (7),

$$\bar{P}_m = \frac{\bar{K}}{\sum_t \bar{K}_t} \tag{7}$$

Here, $\bar{P}_{count}$ represents the passenger's count, $\bar{K}$ represents the number of passengers arriving in the station in a specified time. Similarly, number of passengers in vehicle based on the passenger count is computed by the subsequent condition (8).

$$V_m = V_{m-1}\left(1 - \bar{P}_m\right) + \frac{R_m \bar{K}}{N} \tag{8}$$

Here, N represents the average number of trips in a specified time, R_m is computed by the subsequent condition (9),

$$R_m = \begin{cases} \dfrac{M-m}{M-1} & for\ direction: m \to M \\ \dfrac{m-1}{M-1} & for\ direction: m \to 1 \end{cases} \tag{9}$$

Here, m and M represents the two consecutive stations.

c) Route Distance: Route length represents the sum of the distance among all the mid stations. The route distance is estimated by calculating the distance among each stop from source and destination. The distance between two adjacent nodes are computed by the subsequent condition (10),

$$\bar{D}_{route} = a\frac{\bar{S}_j \bar{D}_k}{dist_{j,k}} \tag{10}$$

Here, $\bar{D}_{route}$ represents the route distance, a signifies the adjustment factor, $\bar{S}_j$ represents the source point of trip at j,

$\overline{D}_k$ represents the destination point trip at k, and $dist_{j,k}$ signifies the distance between the two station points.

d) Average Travel Time: The travel time is the starting and stopping time of vehicle from source to destination. The travel time of passenger is varies based on the different stations. The average travel time is computed by the various transport stations. The travel time of passengers are computed by the subsequent condition (11),

$$\overline{T}_{avg} = \sum_{i=1 to n} TS_{\kappa} \tag{11}$$

Here, $\overline{T}_{avg}$ represents the average travel time, TS represents the travel time between two stations. The travel time is increases based on the increasing distance.

2) Optimal routing in transport network using adaptive objective based beetle swarm optimization (ABSO)

Initially, starting and target points are initialized in the transport network. Moreover, the general attributes of beetle swarm optimization approach is initialized. The beetle swarm optimization approach is the combination of beetle foraging mechanism and swarm optimization. Two beetle antennae of the beetles are used to explore the nearby regions. If the one side antenna finds the more valuable data, then the beetle will move to that antenna side. This optimization approach is based on the behaviour of beetles. The flow diagram of the presented ABSO process is depicted in Fig. 2.

Initially, the input optimization parameters are initialized and the fitness is evaluated based on the average weight of four objectives by utilizing the condition (17). Then the swarm attributes of the optimization algorithm is updated. If the maximum iteration is reached then the optimal global best solution is considered as the optimal routing.

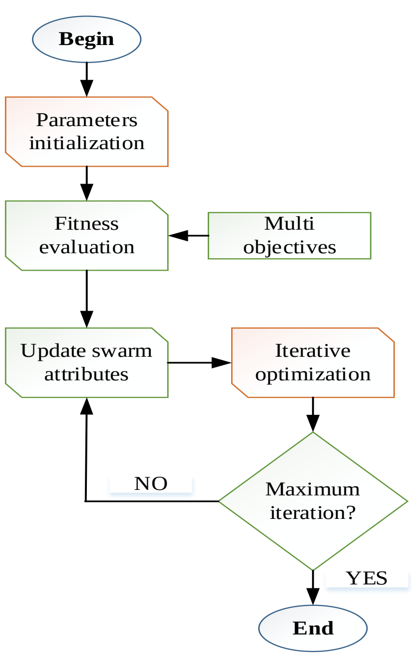

Fig. 2. Flow Diagram of ABSO Routing.

a) Parameters Initialization: The input population of optimization approach is expressed as $\overline{Y} = \left(\overline{Y}_1, \overline{Y}_2, \overline{Y}_3 .....\overline{Y}_m\right)$. Here, m represents the population size. The position $(\overline{P}_a)$ and speed attribute $(\overline{S}_a)$ of beetle is first arbitrarily initialized in a m dimensional search space. Speed attributes of beetle k is expressed as $\overline{V}_k = (v_{k1}, v_{k2}, v_{k3}....v_{kS})^T$. The set of individual best is expressed as $b_i = (b_{i1}, b_{i2}, b_{i3}....b_{iS})^T$. The global best set signified as $\overline{b}_g = (b_{g1}, b_{g2}, b_{g3}....b_{gS})^T$. The speed update of the optimization approach is described in the subsequent condition (12),

$$y_{js}^{t+1} = y_{js}^{t} + \overline{\lambda} v_{js}^{t} + \left(1 - \overline{\lambda}\right) \in_{js}^{t} \tag{12}$$

Here, $$\in_{js}^{t+1} = \partial^t \times v_{js}^t \times F(\overline{Y}) \tag{13}$$

$$y_{rs}^{t+1} = y_{rs}^{t} + v_{is}^{t} \times \frac{\overline{d}^{\,t}}{2} \tag{14}$$

$$y_{ls}^{t+1} = y_{ls}^{t} - v_{ls}^{t} \times \frac{\overline{d}^{\,t}}{2} \tag{15}$$

The position attribute of every beetle individual is updated by the subsequent condition (16),

$$v_{js}^{t+1} = \overline{\varpi} v_{js}^{t} + \overline{c}_1 \overline{r}_1 \left(p_{is}^{t} - \overline{y}_{is}^{t}\right) + \overline{c}_2 \overline{r}_2 \left(p_{gs}^{t} - \overline{y}_{gs}^{t}\right) \tag{16}$$

Here, $s = 1,2,....S$ and $j = 1,2,....m$, t represents the iteration time, $\in_{js}$ signifies the displacement increment examined by the strength of data predicted by the beetle antennae, $\overline{\lambda}$ signifies the decrement factor, $\overline{\varpi}$ signifies the inertia weight, $\overline{r}_1$ and $\overline{r}_2$ are arbitrary function with value ranging from 0 to 1, $\overline{c}_1$ and $\overline{c}_2$ signifies the degree of impact of individual and global best on the beetle, ∂^t signifies the searching step size of beetle, $\overline{d}$ represents the predicted distance of antennae, $F(\overline{Y})$ signifies the fitness function, y_{ls}^{t+1} signifies the predicted distance of left antennae, y_{rs}^{t+1} signifies the predicted distance of right antennae.

b) Fitness Evaluation: Fitness function is evaluated based on four different objectives such as, passenger's count, comfort level of passengers, route distance and average travel time. Fitness of optimal route generation using the proposed optimization approach is described in the subsequent condition (17),

$$F(\overline{Y}) = \min\left(\frac{\overline{P}_m + \overline{D}_{route} + \overline{C}_{\alpha} + \overline{T}_{avg}}{4}\right) \tag{17}$$

Here, $\overline{P}_m$ represents the passenger count computed by the condition (6), $\overline{D}_{route}$ represents the route distance computed by the condition (10), $\overline{T}_{avg}$ represents the average travel time computed by the condition (11), $\overline{C}_\alpha$ represents the comfort level of passengers computed by the condition (9), $F(\overline{Y})$ represents the fitness function for the evaluation of optimal route. The minimum values of fitness function are considered as the global best value ($\overline{b}_g$).

c) Update of Optimization Attributes in Beetle Swarm Optimization: The position and speed attributes of every beetle swarm is updated by the conditions in (12) to condition (16). Further, the fitness function is computed for each updated beetle individuals. Also, the individual best ($\overline{b}_i$) solutions are updated and attains the global ($\overline{b}_g$) solution.

d) Iterative Optimization: The variables ∂^t and $\overline{d}$ of the optimization is updated by the subsequent conditions (18) and (19).

$$\partial^t = 0.95\partial^{t-1} \tag{18}$$

$$\overline{d}^{\,t} = 0.95\overline{d}^{\,t-1} + 0.01 \tag{19}$$

Here, $\overline{d}$ represents the predicted distance of antennae, and ∂^t signifies the searching step size of beetle. Iterations are performed to the maximum by updating these variables to end optimization. Finally, optimal demand based routing is obtained according to the global best output.

IV. Results and Discussion

The experimental results of the proposed demand based on the optimal route generation in transport system is implemented in the working platform of MATLAB. The dataset used for the implementation is Surat city historical transport data. The performance of the presented technique is analyzed with the existing Inverse distance weighting (IDW), and Empirical Bayesian Kriging (EBK) [27], XGBoost, gradient boosting regression model (GDBR), support vector machine regression (SVR), and multiple linear regression model (MLR) [28], agent based scheduled routing [29] techniques in terms of Root mean square error, mean square error, mean absolute error, correlation coefficient, accuracy, vehicle occupancy. The performance metrics utilized for analyzing the presented work is described in the subsequent sections.

A. Dataset Description: Surat City Historical Data

Surat city transportation system historical data is considered for the evaluation of optimal demand based routing performance in the presented work. The passenger detail for this prediction process is taken from Surat city dataset. The passenger details are taken for one year (June 2017 to June 2018) for passenger flow prediction in transportation. The passenger details are taken for one year (June 2017 to June 2018) for optimal route prediction in transportation. The presented DFCM clustering approach output is depicted in Fig. 3.

In Fig. 3, the proposed DFCM clustering output of passenger's count is provided. This provides the accurate details of passengers in every station. The number of passengers in each station is gathered by using this clustering process. Each cluster provides the passenger details of each station. Based on the predicted passenger's count, each route passenger's flow for august month is depicted in Fig. 4.

In Fig. 4, august month data of Surat city transportation dataset is examined for evaluating the Passenger's count variations of route 1. The passenger count of route 1 is depicted in Fig. 4 for each day in a month. Here, RAST, SAHD, KADB, UDHD and ADGS represent the different stations in the considered route 1. The passenger count of each day is calculated by adding the passenger details of each trip of the day. The passenger count in RAST station is higher than the other stations. Similarly, passenger's count variations of route 2 are analyzed with august month data and it are depicted in Fig. 5.

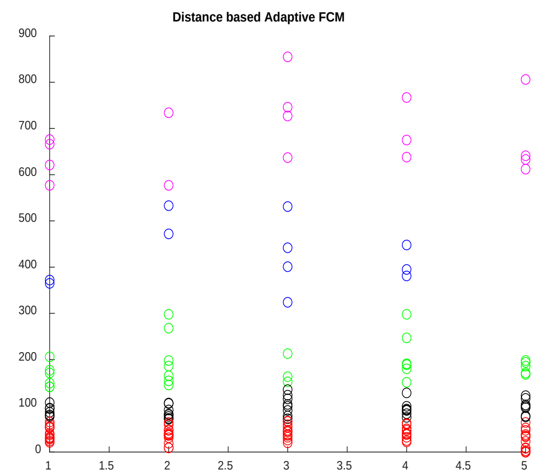

Fig. 3. DFCM Clustering Output.

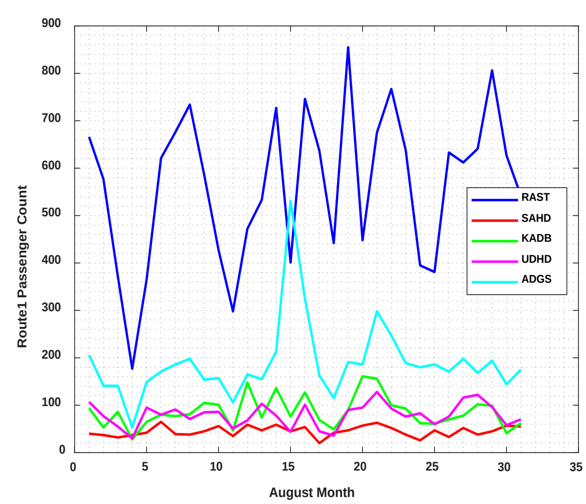

Fig. 4. Passenger's Count Variations of August Month with Route 1.

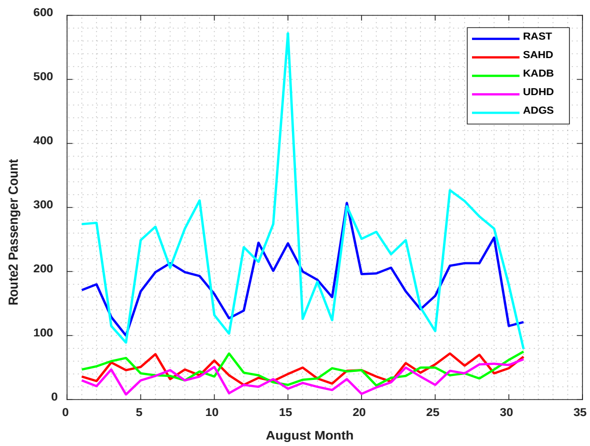

Fig. 5. Passenger's Count Variations of August Month with Route 2.

In Fig. 5, Passenger's count variations of route 2 are examined for evaluating the august month data of Surat city transportation dataset. In Fig. 5, the passenger count of each station of route 2 is depicted. Here, RAST, SAHD, KADB, UDHD and ADGS characterize the different stations of the considered route 2. The passenger counts in route 2 with different stations are provided. This demonstrates that the passenger count in station name ADGS is higher than other stations. Similarly, the daily passenger flow is analyzed for different stations with various trip timings are depicted in Fig. 6.

In Fig. 6, daily passenger's flow of different stations such as, 10, 1003, 1015, 2040, and 2799 is analyzed with various bus timings from morning 6 O' clock to night 9 O'clock. This demonstrates that the daily passenger flow of the station number 10 is higher than the other stations. Moreover, the passenger count is at peak at 9 o clock timing than other timings. Here, the daily passenger flow of each trip timing details is demonstrated for a month. The passenger count of each trip for one month is illustrated for various stations. Moreover, the passenger's demand for one month data is depicted in Fig. 7.

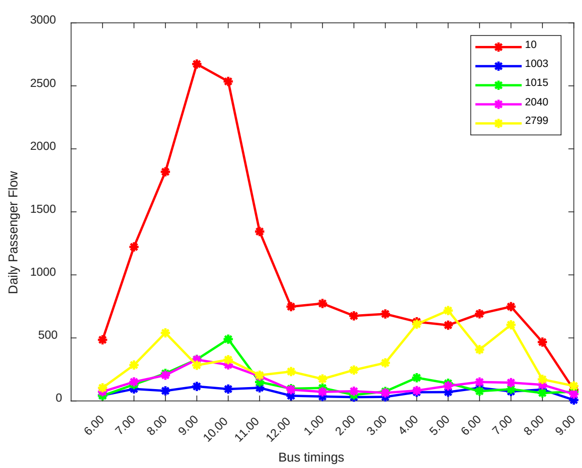

Fig. 6. Daily Passenger's Flow of different Stations with Various Trip Timings.

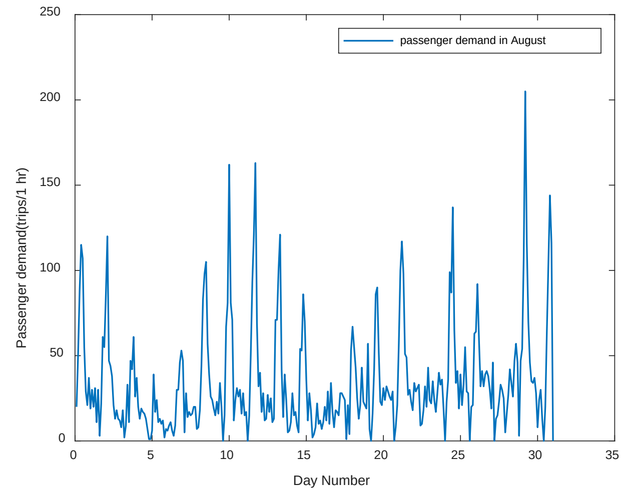

Fig. 7. Passenger Demand in August Month.

The Fig. 7 depicts that the demand of passengers for a month. Based on the passengers demand, transport needs to be allocated with optimal routing approach. Here, the passenger demand is calculated based on the passenger data for 30 days. Furthermore, routing is performed based on the adaptive beetle swarm optimization approach for august month data. The effective objectives such as passenger count, travel time, comfort level and distance are considered for re-routing. In this way, an optimal demand based routing is generated to enhance the quality of service. Moreover, the passenger demand for a week based data for a month is depicted in Fig. 8.

In Fig. 8, passenger demand is evaluated for an august month data. Here, the passenger demand is evaluated for each day in a week for a month. This figure provides the passenger demand data for a month on weekly basis. The demand based objectives are considered for each pair of stops in the transportation system. The demand based objective measure for the pair of stations 1 and station 5 is depicted in Fig. 9.

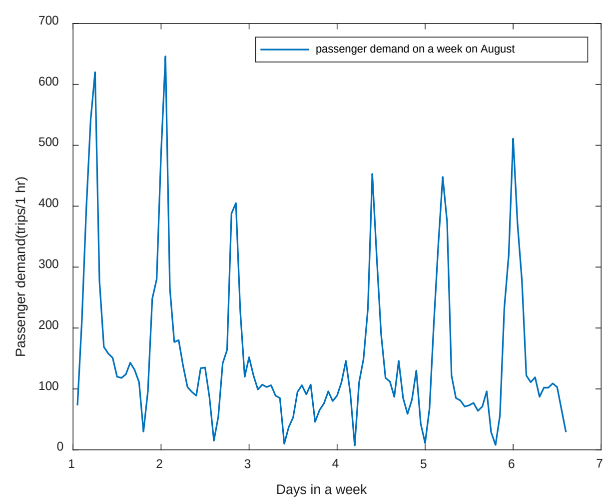

Fig. 8. Passenger Demand of Week based Data for an August Month.

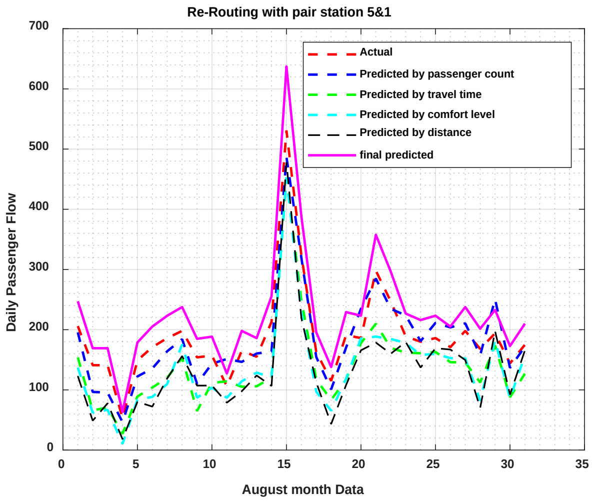

Fig. 9. Demand based Routing with Pair of Station 1 & 5.

In Fig. 9, august month data is considered to predict the demand based routing. Here, the daily passenger flow is illustrated by considering various parameters like passenger count data, travel time, comfort level and distance. Finally, daily passenger is depicted by considering all these parameters. Here, pair of stations 1 and station 5 is illustrated to prove the effectiveness of proposed demand based routing. The proposed route 2 is obtained as an optimal routing. Similarly, optimal routing based on the effective objectives with pair of station 3 and station 4 is depicted in Fig. 10.

In Fig. 10, demand based routing with pair of station 3 and station 4 is provided. Here, august month data is considered to analyze the routing based on different objectives. The combined objectives output is providing higher performance than the single objectives. This illustration proved that the effectiveness of proposed demand based routing. Similarly, optimal routing based on the effective objectives with pair of station 1 and station 2 is depicted in Fig. 11.

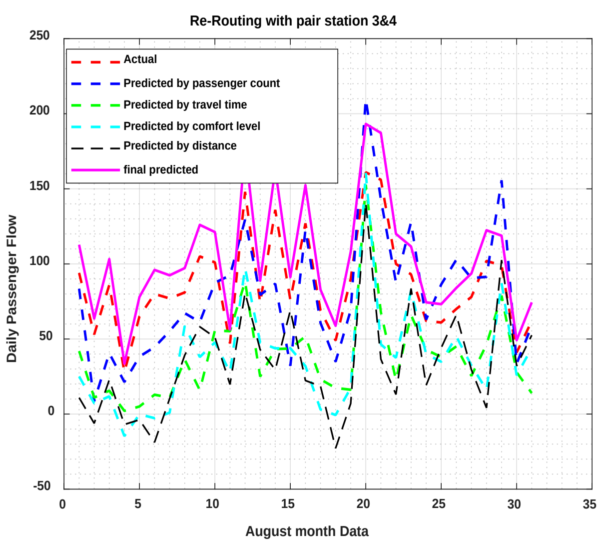

Fig. 10. Demand based Routing with Pair of Station 3 & 4.

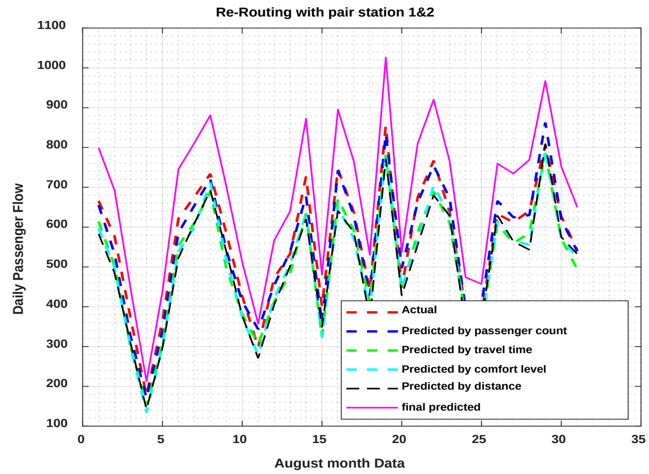

Fig. 11. Demand based Routing with Pair of Station 1 & 2.

In Fig. 11, optimal routing is evaluated with pair of station 1 and station 2. Here, the prediction performance is analyzed without objectives and with objectives such as passenger count, travel time, comfort level, distance. The final predicted output is based on all the combined objectives. Optimal routing based on the effective demand based objectives for the pair of station 4 and station 5 is depicted in Fig. 12.

In Fig. 12, routing performance is evaluated with the pair of stations 4 and station 5. This proves that the proposed demand based optimal routing achieves enhanced performance than the single objective based predictions. The passenger's flow with the proposed demand based routing is enhanced. This enhancement proves that the service quality of the transportation system.

B. Performance Metrics

In this section, different performance metrics such as Root mean square error, mean square error, mean absolute error, correlation coefficient, accuracy, vehicle occupancy are described in the subsequent sub sections.

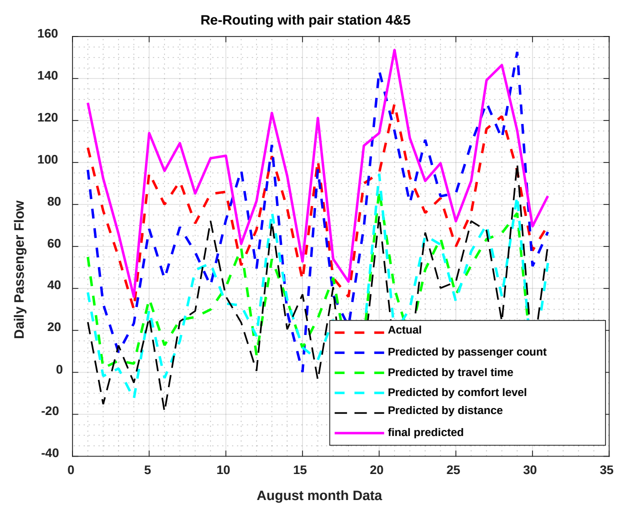

Fig. 12. Demand based Routing with Pair of Station 4 & 5.

1) Correlation coefficient: Correlation coefficient is evaluating the correlation among the considered variables. The correlation coefficient range is 0 to 1. The performance is proved by obtaining the value closer to 1. The correlation is estimated between the attained optimal route and the other existing routes. It is computed by the subsequent condition (20),

$$\overline{C}_C = \frac{Cov(y_k, z_k)}{\sqrt{Var[y_k] - Var[z_k]}} \tag{20}$$

Here, y_k and z_k represents the predicted optimal and original routing, $Cov(y_k, z_k)$ represents the covariance of y_k and z_k, $Var[y_k]$ represents the variance of y_k and $Var[z_k]$ represents the variance of z_k. In the optimal routing prediction, correlation coefficient is used to examine the linear correlation among the predicted optimal data and original data.

2) Root means squared error (RMSE): This RMSE measure is used to examine the variation between predicted and the original values. This demonstrates the model accuracy from the perception of predicted deviation. This is computed by the subsequent condition (21).

$$\overline{R}_{MSE} = \sqrt{\frac{\sum_{k=1}^{m} |y_k - z_k|^2}{m}} \tag{21}$$

Here, $\overline{R}_{MSE}$ represents the root mean squared error. This measure is used to prove the prediction accuracy rate by minimizing the error in optimal routing.

3) Mean absolute error (MAE): The mean absolute error performance is the average of absolute values of the variation among all the predicted values and the original values. This accurately predicts the error of the predicted output. The MAE measure is computed by the subsequent condition (22).

$$\overline{M}_{AE} = \frac{1}{m} \sum_{k=1}^{m} |y_k - z_k| \tag{22}$$

Here, $\overline{M}_{AE}$ represents the mean absolute error. The minimum value of mean absolute error proves that the proposed demand based optimal routing is effective technique demand based optimal route prediction.

4) Mean squared error (MSE): This measure is similar to the MAE performance metric but only the difference is taking square in the mean differences of the input and predicted data. It is calculated with the mean of the square difference among the original input and the predicted data. It is evaluated with the subsequent condition (23).

$$\overline{MsE} = \frac{1}{T} \sum_{k=1....T} (z_k - \tilde{z}_k)^2 \tag{23}$$

Here, $\overline{MsE}$ represents the MSE, T represents the total considered data, z_k and $\tilde{z}_k$ denotes the original data and predicted data correspondingly.

5) Accuracy: Accuracy measure is the ratio of correctly predicted data to the total number of observations. Moreover, it is the percentage of accurateness for predicting the data. It is evaluated utilizing the subsequent condition (24).

$$\overline{A}_y = \frac{\overline{c}_P}{T_p} \tag{24}$$

Here, $\overline{A}_y$ represents the accuracy, $\overline{c}_P$ represents the correct predictions, and T_p represents the total predictions.

6) Vehicle occupancy: Vehicle occupancy is determined by the occupancy rate and it is equivalent to average number of passengers in the vehicle. It is computed based on the average numbers of passengers used the vehicles in the same route. The average computation is attained from the number of trips in the same route for a day. The passenger's count of each trip is estimated to attain the vehicle capacity. The vehicle occupancy is computed by the subsequent condition (25).

$$V_O = \frac{\sum T_N (P_C)}{N} \tag{25}$$

Here, V_O represents the vehicle occupancy, T_N represents the number of vehicle trips, N represents the Trips number, and P_C represents the passengers in each trip.

C. Performance Analysis

In this section, performance of the presented approach is examined with the existing strategies in terms of Root mean square error, mean square error, mean absolute error, correlation coefficient, accuracy, vehicle occupancy. The performance metrics and their validations are examined with the various existing approaches. The RMSE performance evaluation is analyzed with the existing approaches for each station is depicted in Fig. 13.

In Fig. 13, the RMSE performance demonstration is provided with existing approaches for each station. The RMSE value of the presented approach for each station is station 1 (0.2629), station 2 (0.2217), station 3 (0.2474), station 4 (0.2629), and station 5 (0.0768) respectively. The presented demand based route prediction is examined with the existing XGBoost, gradient boosting regression model (GDBR), support vector machine regression (SVR), and multiple linear regression model (MLR) approaches. Here, the RMSE performance is evaluated for 5 stations with the existing approaches. The RMSE performance on varying stations is portrayed in Table I.

In Table I, RMSE measure values are lesser than the different existing approaches. The obtained RMSE of presented approach for each station is station 1 (0.2629), station 2 (0.2217), station 3 (0.2474), station 4 (0.2629), and station 5 (0.0768) correspondingly. The performance analysis of the proposed demand based routing with existing approaches is depicted in Fig. 14.

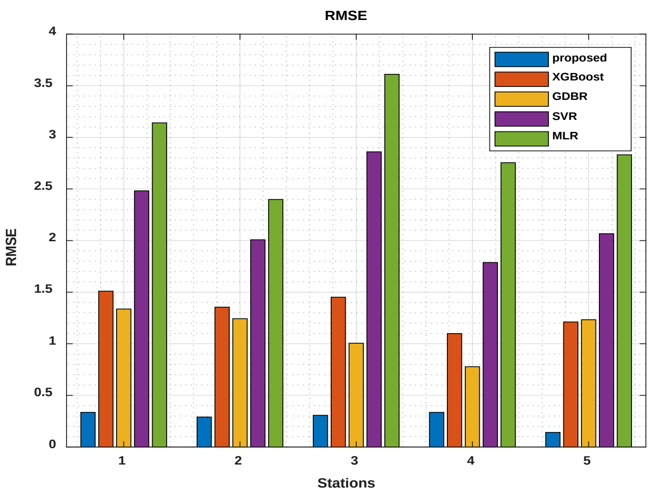

Fig. 13. Performance Analysis on RMSE for Each Station.

TABLE I. RMSE PERFORMANCE ON DIFFERENT STATIONS

Technique	XGBoost	GDBR	SVR	MLR	Proposed
STATION 1	1.1043	1.1131	2.5570	3.0657	0.2629
STATION 2	1.4011	1.0558	1.6021	3.1922	0.2217
STATION 3	0.8773	1.3297	1.3851	3.0004	0.2474
STATION 4	0.6586	0.9369	3.0113	2.7080	0.2629
STATION 5	1.0901	1.5924	2.7080	3.5162	0.0768

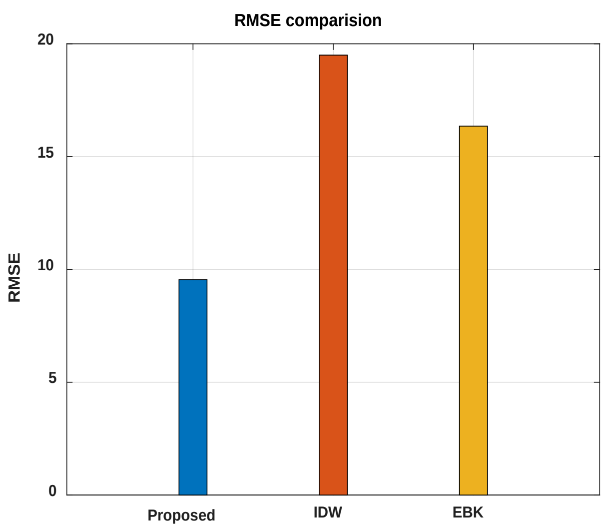

Fig. 14. Performance Analysis on RMSE.

In Fig. 14, RMSE performance evaluation is provided with existing Inverse distance weighting (IDW), and Empirical Bayesian Kriging (EBK) [27] approaches. The existing EBK, and IDW approaches are used Orissa dataset, which is overcome by the proposed approach with surat city dataset. Here, the RMSE value of presented approach is 9.5. This error value is much lesser than the existing IDW (19.4994) and EBK (16.3527) approaches. This demonstrates that the presented approach provides reduced RMSE than the existing approaches to prove the efficiency of the proposed approach. The comparison analysis on RMSE is mentioned in Table II.

In Table II, the RMSE of the presented approach is compared with the existing approaches. The proposed approach attains reduced RMSE than the existing IDW, EBK [27] approaches. The presented approach attains enhanced performance than the existing techniques by reducing the error rate. The comparison analysis of MSE is examined with the existing approaches are depicted in Fig. 15.

In Fig. 15, performance of the presented approach in terms of MSE is analyzed with the different existing approaches like XGBoost, GDBR, SVR, and MLR. The MSE value of presented approach for each station is station 1 (0.3007), station 2 (0.2548), station 3 (0.2500), station 4 (0.3007), and station 5 (0.2544) respectively. This clearly demonstrates that the presented demand based optimal routing approach provides enhanced performance with less error. The performance comparison on MSE is mentioned in Table III.

TABLE II. COMPARISON ANALYSIS ON RMSE

Techniques	Dataset	Root Mean squared error (%)
Inverse distance weighting	Orissa dataset	19.4994
Empirical Bayesian Kriging	Orissa dataset	16.3527
Proposed	Surat city historical dataset	9.5

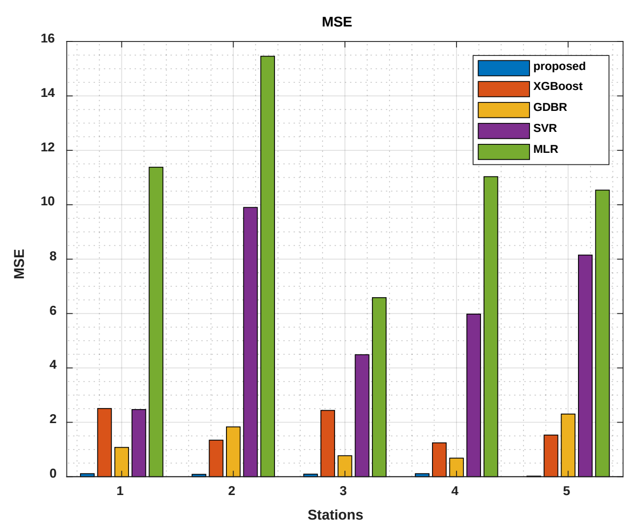

Fig. 15. Performance Analysis on MSE.

In Table III, the performance comparison on MSE is provided. Here, the performance of the presented approach is compared with the different existing approaches like GDBR, SVR, XGBoost, and MLR. The attained MSE value of presented approach for each station is station 1 (0.3007), station 2 (0.2548), station 3 (0.2500), station 4 (0.3007), and station 5 (0.2544) correspondingly. Moreover, the comparison on mean absolute error is depicted in Fig. 16.

In Fig. 16, mean absolute error performance of the proposed and the existing strategies for each station is provided. This proves that the presented demand based routing provides enhanced performance than the existing XGBoost, GDBR, SVR, and MLR approaches. The attained MAE value of presented approach for each station is station 1 (0.2007), station 2 (0.3548), station 3 (0.2500), station 4 (0.5007), and station 5 (0.3544) correspondingly. The performance on MAE for varying stations is portrayed in Table IV.

TABLE III. MSE PERFORMANCE ON DIFFERENT STATIONS

Technique	XGBoost	GDBR	SVR	MLR	Proposed
STATION 1	1.2195	1.2390	6.5384	9.3988	0.3007
STATION 2	1.9630	1.1147	2.5669	10.1900	0.2548
STATION 3	0.7697	1.7681	1.9185	9.0023	0.2500
STATION 4	0.4338	0.8778	9.0681	7.3333	0.3007
STATION 5	1.1884	2.5357	4.9333	12.3639	0.2544

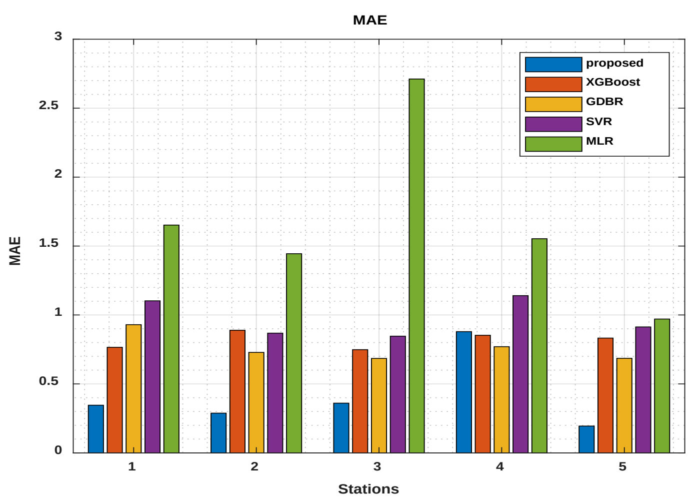

Fig. 16. Performance Analysis on MAE (%).

TABLE IV. MAE PERFORMANCE ON DIFFERENT STATIONS

Technique	XGBo-ost	GDBR	SVR	MLR	Proposed
STATION 1	0.7415	0.5089	0.8601	2.0907	0.2007
STATION 2	0.7542	0.6676	0.8905	2.3933	0.3548
STATION 3	0.8470	0.5624	1.2911	1.4605	0.2500
STATION 4	0.8975	0.8437	1.4642	0.9596	0.5007
STATION 5	0.8760	0.5943	0.9015	1.2081	0.3544

The Table IV proves that the presented methodology provides enhanced performance in terms of MAE than the XGBoost, GDBR, SVR, and MLR approaches. Moreover, the performance efficiency is examined by the vehicle occupancy is depicted in Fig. 17.

In Fig. 17, vehicle occupancy performance is demonstrated with the existing agent based scheduled routing [29]. Here, the vehicle occupancy is evaluated based on the number of passengers used the vehicle for a month. This illustrates that the presented routing process increases the usage of passengers. This proved that the performance of demand based optimal routing is effective than the other transportation routings. The vehicle occupancy is enhanced by the better quality service of the transportation based on the presented methodology. The performance on accuracy for varying stations is portrayed in Table V.

The Table V proves that the presented methodology provides enhanced performance in terms of accuracy than the deep neural network (DNN), Random forest (RF), Linear regression (LR) approaches. The attained accuracy of presented approach for each station is station 1 (99.51%), station 2 (98.14%), station 3 (98.69%), station 4 (97.61%), and station 5 (98.19%) correspondingly. The accuracy performance of the proposed optimal routing is depicted in Fig. 18.

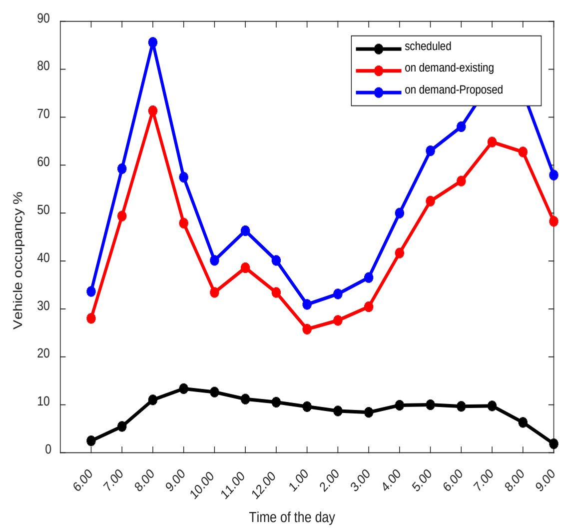

Fig. 17. Performance Analysis on Vehicle Occupancy (%).

TABLE V. ACCURACY PERFORMANCE ON DIFFERENT STATIONS

Technique	DNN (%)	LR (%)	RF (%)	Proposed (%)
STATION 1	98.26	94.52	94.08	99.51
STATION 2	98.58	94.52	94.08	98.14
STATION 3	98.5858	94.52	94.088	98.69
STATION 4	98.58	94.52	94.087	97.61
STATION 5	98.58	94.52	94.089	98.19

In Fig. 18, the presented approach is analyzed with the existing approaches in terms of accuracy performance. This proved that the presented approach provides enhanced accuracy performance than the existing strategies. The performance on correlation coefficient for varying stations is mentioned in Table VI.

The Table VI proves that the presented methodology provides enhanced performance in terms of correlation coefficient than the XGBoost, GDBR, SVR, and MLR approaches. The attained correlation coefficient of presented approach for each station is station 1 (0.8993) station 2 (0.1000), station 3 (0.2800), station 4 (0.2000), and station 5 (0.4800) correspondingly. The comparison analysis in terms of correlation coefficient is depicted in Fig. 19.

In Fig. 19, the correlation coefficient performance is examined with the different existing approaches. The correlation coefficient of presented approach for each station is station 1 (99.51%), station 2 (98.14%), station 3 (98.69%), station 4 (97.61%), and station 5 (98.19%) correspondingly. Here, the proposed approach achieved a high correlation coefficient for each station than the existing the XGBoost, GDBR, SVR, and MLR approaches. Here, this validation is evaluated for each station in the considered Surat city transportation dataset. This proposed methodology provides enhanced results in demand based optimal routing. Furthermore, this methodology enhances the quality of transportation service.

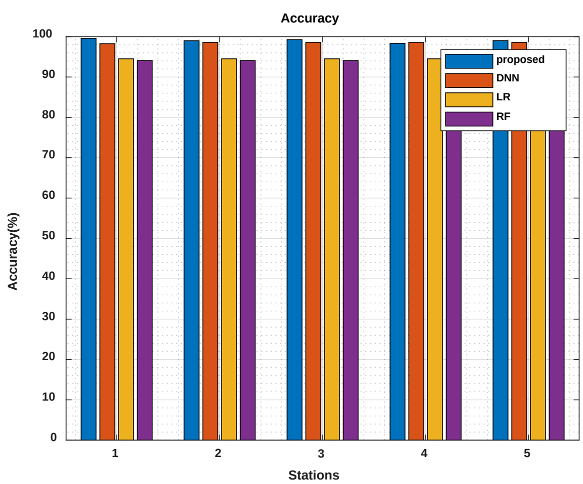

Fig. 18. Performance Analysis on Accuracy (%).

TABLE VI. CORRELATION COEFFICIENT PERFORMANCE ON DIFFERENT STATIONS

Technique	XGBoost	GDBR	SVR	MLR	Proposed
STATION 1	0.6259	0.6529	0.5598	0.2788	0.8993
STATION 2	0.5876	0.3125	0.5281	0.1164	0.1000
STATION 3	0.6263	0.4114	0.5977	0.2863	0.2800
STATION 4	0.4736	0.6480	0.2810	0.2754	0.2000
STATION 5	0.5624	0.5944	0.4928	0.1643	0.4800

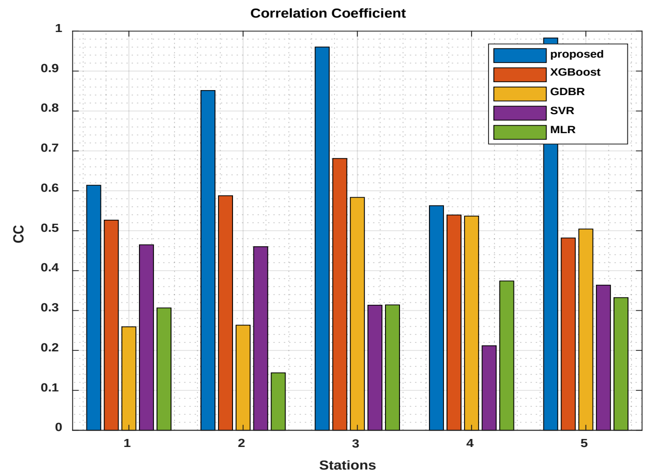

Fig. 19. Performance Analysis on Correlation Coefficient.

V. CONCLUSION

This paper presented a demand based effective routing in transportation network. At first, Distance based adaptive Fuzzy C-means clustering approach (DFCM) is utilized for clustering the passengers in each stop. Here, the number of cluster members in each cluster is equivalent to the passenger's count of every station. Afterwards, adaptive beetle swarm optimization approach based routing is performed based on the clustered data. Further, re-routing is performed based on the effective objectives such as passenger's count, comfort level of passengers, route distance and average travel time using ABSO approach. In the optimal route generation approach, objectives are analysed in each pair of stations in the route using the ABSO approach. This ABSO approach enhances transportation system by providing demand based routing. The presented work is examined with the various existing approaches in terms of different performances like RMSE (9.5%), Mean error (0.254%), vehicle occupancy (85%), MAE (0.3007%), correlation coefficient (0.8993), and accuracy (99.57%). This proved that the presented approach provides better performance than the existing routing approaches in transportation system. In future, this demand based transportation is further enhanced with improved deep learning based approaches to decrease the time consumption. Increasing number of objectives will further improve the performance of demand based routing. Moreover, the demand based transportation will be examined with other larger datasets.

ACKNOWLEDGMENT

I sincerely thank Nirbhay Chaubey for their guidance and encouragement in carrying out this research work.

REFERENCES

[1] M. Nassereddine, H. Eskandari, "An integrated MCDM approach to evaluate public transportation systems in Tehran, Transportation Research Part A: Policy and Practice," vol. 106, pp. 427-439, 2017.

[2] X. Wan, H. Ghazzai, Y. Massoud, "Mobile Crowdsourcing for Intelligent Transportation Systems: Real-Time Navigation in Urban Areas," IEEE Access, vol. 7, pp. 136995-137009, 2019.

[3] T. Afrin, N. Yodo, "A Survey of Road Traffic Congestion Measures towards a Sustainable and Resilient Transportation System," Sustainability, vol. 12, pp. 4660, (2020).

[4] H. Xiong, A. Vahedian, X. Zhou, Y. Li, J. Luo, "Predicting traffic congestion propagation patterns: A propagation graph approach," In Proceedings of the 11th ACM SIGSPATIAL International Workshop on Computational Transportation Science, pp. 60-69, 2018.

[5] H. Xiong, A. Vahedian, X. Zhou, Y. Li, J. Luo, "Predicting traffic congestion propagation patterns: A propagation graph approach," In Proceedings of the 11th ACM SIGSPATIAL International Workshop on Computational Transportation Science, pp. 60-69, 2018.

[6] J. Weng, X. Di, C. Wang, J. Wang, L. Mao, "A bus service evaluation method from passenger's perspective based on satisfaction surveys: A case study of Beijing," China, Sustainability, vol. 10, no. 8, pp. 2723, 2018.

[7] S. Güner, "Measuring the quality of public transportation systems and ranking the bus transit routes using multi-criteria decision making techniques," Case studies on transport policy, vol. 6, no. 2, pp. 214-224, 2018.

[8] M. Javid, S. Hussain, M. Anwaar, "Passenger's Perceptions on Prospects of Qingqi Paratransit Public Transport Service in Lahore," Iranian Journal of Science and Technology, Transactions of Civil Engineering, vol. 44, pp. 185-195, 2020.

[9] Y. Wang, Z. Zhang, M. Zhu, H. Wang, "The Impact of Service Quality and Customer Satisfaction on Reuse Intention in Urban Rail Transit in Tianjin," China. SAGE Open, vol. 10, pp. 215824401989880, 2020.

[10] A. Alkharabsheh, S. Moslem, S. Duleba, "Evaluating Passenger Demand for Development of the Urban Transport System by an AHP Model with the Real-World Application of Amman," Applied Sciences, vol. 9, pp. 4759, 2019.

[11] Ö. Bilişik, Ş. Şeker, N. Aydın, N. Güngör, H. Baraçlı, "Passenger Satisfaction Evaluation of Public Transportation in Istanbul by Using Fuzzy Quality Function Deployment Methodology," Arabian Journal for Science and Engineering, vol. 44, pp. 2811-2824, 2019.

[12] V. Cacchiani, J. Qi, L. Yang, "Robust optimization models for integrated train stop planning and timetabling with passenger demand uncertainty," Transportation Research Part B: Methodological, vol. 136, pp. 1-29, 2020.

[13] Ş. İmre, D. Çelebi, Measuring Comfort in Public Transport: A case study for İstanbul, Transportation Research Procedia, vol. 25, pp. 2441-2449, 2017.

[14] C. Zhang, Y. Liu, W. Lu, G. Xiao, "Evaluating passenger satisfaction index based on PLS-SEM model: Evidence from Chinese public transport service," Transportation Research Part A: Policy and Practice, vol.120, pp. 149-164, 2019.

[15] X. Shen, S. Feng, Z. Li, B. Hu, "Analysis of bus passenger comfort perception based on passenger load factor and in-vehicle time," SpringerPlus, vol. 5, 2016.

[16] F. Kurtulmuşoğlu, F. Pakdil, K. Atalay, "Quality improvement strategies of highway bus service based on a fuzzy quality function deployment approach," Transportmetrica A: Transport Science, vol. 12, pp. 175-202, 2016.

[17] S. Ordonez Medina, B. Wang, "Public transport routing including fixed schedule, shared on-demand and door-to-door services," Procedia Computer Science, vol. 151, pp. 846-851, 2019.

[18] I. Katsaragakis, I. Tassopoulos, G. Beligiannis, "Solving the Urban Transit Routing Problem Using a Cat Swarm Optimization-Based Algorithm," Algorithms, vol. 13, pp. 223, 2020.

[19] T. Vlachopanagiotis, K. Grizos, G. Georgiadis, I. Politis, "Public Transportation Network Design and Frequency Setting: Pareto Optimality through Alternating-Objective Genetic Algorithms," Future Transportation, vol. 1, pp. 248-267, 2021.

[20] M. Owais, M. Osman, "Complete hierarchical multi-objective genetic algorithm for transit network design problem," Expert Systems with Applications, vol. 114, pp. 143-154, 2018.

[21] M. Owais, M. Osman, "Complete hierarchical multi-objective genetic algorithm for transit network design problem," Expert Systems with Applications, vol. 114, pp. 143-154, 2018.

[22] O. Cosma, P. Pop, D. Dănciulescu, "A novel matheuristic approach for a two-stage transportation problem with fixed costs associated to the routes," Computers & Operations Research, vol. 118, pp. 104906, 2020.

[23] J. Liu, P. Mirchandani, X. Zhou, "Integrated vehicle assignment and routing for system-optimal shared mobility planning with endogenous road congestion," Transportation Research Part C: Emerging Technologies, vol. 117, pp. 102675, 2020.

[24] J. Narayan, O. Cats, N. van Oort, S. Hoogendoorn, "Integrated route choice and assignment model for fixed and flexible public transport systems," Transportation Research Part C: Emerging Technologies, vol. 115, pp. 102631, 2020.

[25] M. Owais, A. Alshehri, "Pareto Optimal Path Generation Algorithm in Stochastic Transportation Networks," IEEE Access, vol. 8, pp. 58970-58981, 2020.

[26] P. Heyken Soares, "Zone-based public transport route optimisation in an urban network," Public Transport, vol. 13, pp. 197-231, 2020.

[27] K.K.R. Samal, K.S. Babu, S.K. Das, "ORS: The optimal routing solution for smart city users," In Electronic systems and intelligent computing, Springer, Singapore, pp. 177-186, 2020.

[28] X. Kong, M. Li, T. Tang, K. Tian, L. Moreira-Matias, F. Xia, "Shared Subway Shuttle Bus Route Planning Based on Transport Data Analytics," IEEE Transactions on Automation Science and Engineering, vol. 15, pp. 1507-1520, 2018.

[29] S. Liyanage, H. Dia, "An Agent-Based Simulation Approach for Evaluating the Performance of On-Demand Bus Services," Sustainability, vol. 12, pp. 4117, 2020.