Research Paper Summary

TITLE	AUTHOR	SOURCE	SUMMARY
MODELING AND ANALYSIS OF BUS SCHEDULING SYSTEMS OF URBAN PUBLIC BUS TRANSPORT[1]	EshetieBerhan, DejeneMengistu, BerhanuBeshah and Daniel Kitaw	International Journal of Computer Information Systems and Industrial Management Applications	The objective of this paper is to develop an optimum bus assignment method using Linear Programming (LP). After thorough analysis of the existing bus scheduling system, the LP model has been developed and is used to determine the optimal number of buses for each route in four shifts. The output of the LP-model is then validated with the performances of the existing systems. The findings of their study show that the new model reveals better performances on the operating costs, bus utilization and trips and distance covered compared with the existing scheduling system. The enterprise's bus utilization was improved by the new system and cut costs on the one hand and improves the service quality to passengers on the other hand. Bus scheduling is one of the operations planning process in bus Transport Company that deals with the proper assignment of buses to serve the expected passenger demand. The decision-making process of bus assignment is however a trade-off between service quality and operating cost for the bus operating companies. It is because; using too many buses incurs more operating cost to the company whereas too few buses decrease the service quality level. But, the enterprise uses a fixed number of buses scheduled per route in its operation throughout the day. This resulted in, the fact that, some buses move empty while others are overcrowded, which subsequently result in poor performance on bus utilization, distance travelled, number of trips and service quality.
BUS SCHEDULING MODEL: A LITERATURE REVIEW[2]	Mohammad HesamHafezi, Amiruddin Ismail and Ramez A. Al-Mansob	Regional Engineering Postgraduate Conference (EPC) 2011	This paper mainly talks about the various agencies involved in the public transportation system, mainly in Bus scheduling models. These agents are mainly: passenger, bus authority and traffic. It also specifies performance characteristics of bus services as: reliability, frequency, capacity, safety and costs. In this paper, the authors review some bus scheduling models and show some of the effective parameters to obtain proper bus scheduling models. The Bus scheduling models reviewed in this paper are: • In 2002 a scheduling model based on interrelationship between passenger trip demands and bus trip supplies for inter-city bus carriers have been studied by Yan and Chen . • HWE et al. in 2006 review merging bus route for obtained proper bus scheduling model in central business district.They proposed a merging bus routes method for reduce traffic congestion problem with increasing bus occupancy in central business district. They review overlapping of bus routes and fleet size. Also, they obtained rate of passengers in different operation time during day. They by merging routes which they have more than 60% overlap obtained a new method for bus operation. The proposed method could decrease fleet size and the number of bus stopping activities • Also, van Oudheusden and Zhu in 1995 have proposed a trip frequency scheduling for determination of trip frequencies problem which this way reduced trip frequencyduration different period. They achieved that overcrowding of buses can depend by insufficient planning in addition fleet size and traffic congestion (van Oudheusden and Zhu 1995). • In 2009 Chen et al. analyzed bus operation reliability at the stop, route and network levels. They achieved by increasing distance between a bus stop and the origin terminal, reliability of bus service will be declined. Also,

			bus service reliability greatly decreases when this distance to increase to more than 30 km (Chen et al. 2009) The paper then concludes with some suggestions and inferences like: Generally, delay of buses are due to some cases: more demand of passengers, depletion of fleet size, traffic characteristic and frequency of buses. During peak-hour traffic, disorganization of bus scheduling is higher than non-peak-hour traffic. Using an exclusive bus lane for bus operation can reduce disorganization (Wirasinghe and Vandebona2010). Moreover, to improve bus service in crowed areas, the mixed traffic lanes and exclusive bus lane can be used together. For improving bus scheduling, they suggest some recommendations: if a crowded bus arrives after a long waiting time, indeed there is usually an empty bus behind this bus and the crowded bus should skip few stops in such case.
REAL-TIME OPTIMAL BUS SCHEDULING FOR A CITY USING A DTR MODEL[3]	MakrandWagale, AjitPratap Singh , Ashoke K Sarkar And Srinivas Arkatkar	2nd Conference of Transportation Research Group of India (2nd CTRG)	Timetabling and vehicle scheduling is the basis of security and efficiency for various bus enterprises. It is necessary to take into account the passenger travel demand to meet both the social and economic benefits for these bus enterprises. In this paper, a Demand and Travel time Responsive (DTR) model has been developed to actualize a timetable for each bus stop on the basis of optimal bus frequency. • This paper presents a model to optimize the bus scheduling by taking into consideration both bus stop and route segments of the city in an integrated manner. In this study different real time data event parameters, such as bus stop departures and arrivals for buses operating on a line-based time-table and bus traffic costs have been applied to optimize the bus scheduling process. A bus headway time-table is also being developed. • The proposed model has been verified by taking a case study of Jaipur city. The relevant data have been collected from Jaipur City Transport Services Limited (JCTSL). The sensitivity analysis for various parameters and assumption used in proposed model has been applied to assess the reliability of the optimal solution. The DTR model for an optimal bus scheduling developed herein is based on a holistic, integrated, systems-oriented approach, which clearly demonstrates the overall usage of the model. • The objective for this study was to develop a rational schedule plan that balances mutual trade-offs and is based on scientific analyses using stop-based or segment-based microdata collected through the system. It was found that the concentration of passengers was mainly on two stops which were main departure spots on the entire route and variation of total traffic cost with these stops with others was also high.
BUS TRANSIT SERVICE PLANNING AND OPERATIONS IN A COMPETITIVE ENVIRONMENT[4]	El-Geneidy, A., Hourdos, J., & Horning, J.	Journal of Public Transportation	In this research paper, authors correlate travel time obtained from buses to travel time obtained from floating vehicles, in the Twin Cities Metropolitan region. This research helps in introducing more reliable estimates of travel time for planning new and competitive transit services. Specifically, this work studies two bus routes over a variety of different roadway types and traffic conditions and produced statistical models that can estimate travel time based on measurements collected from buses and regular vehicle probes. The generated models revealed the characteristics causing bus service to be generally slower. Altering bus route characteristics can reduce overall travel time and minimize the travel time disparity between buses and private vehicles. In particular, the models presented in this paper lend support to bus only shoulder policies, stop consolidation, serving major streets with fewer stop signs and implementation of smart transit signal priority.In this research we analyse information from

			different roadway types (freeways, arterials, and local streets) to uncover potential traffic flow related dependencies. In order to establish the relationship between travel times for buses and private vehicles in the study area, each bus trip was matched with a probe vehicle trip that departed at approximately the same time. The paper then conducts various t-tests to establish the relationship between the buses and private vehicles. It also selects appropriate attributes to consider, thus ensuring that the tests only consider contributing attributes. After finding sufficiently strong evidence supporting their hypothesis, the authors then design and define a regression model to predict the time required by the Buses and Private vehicles at any given time of the day. This predictions can help decide when to prefer traveling through private vehicles and when to travel using buses.
OPTIMAL MULTI-VEHICLE TYPE TRANSIT TIMETABLING AND VEHICLE SCHEDULING[5]	Avishai (Avi) Ceder	14th EWGT & 26th MEC & 1st RH	This paper addresses two activities: timetable development and vehicle-scheduling with different vehicles types. Alternative timetables are constructed with either even-headways, but not necessarily with even passenger loads or even average passenger loads, but not even headways. A method to construct timetables with the combination of both even-headway and even-load concepts has been developed for multi-vehicle sizes in this paper. The vehicle-scheduling problem is based on given sets of trips and vehicle types arranged in decreasing order of vehicle cost. This problem can be formulated as a cost-flow network problem with an NP-hard complexity level. Thus, a heuristic algorithm is developed and few examples are used as an expository device to illustrate the procedures developed. This paper defines Timetabling and Vehicle Scheduling as: • Vehicle scheduling refers to the problem of determining the optimal allocation of vehicles to carry out all thetrips in a given transit timetable. A chain of trips is assigned to each vehicle including possible deadheading (DH) orempty trips. The number of feasible solutions to this problem is extremely high, especially in the case in which thevehicles are based in multiple depots. • The problem of finding the best dispatching policy for transit vehicles on fixed routes has a direct impact onconstructing timetables. This dispatching-policy problem, which has been dealt with quite extensively in theliterature, can be categorized into four groups: (1) models for an idealized transit system, (2) simulation models, (3)mathematical programming models, and (4) data-based models. In each step, of the heuristic procedure, buses are assigned departure times based on an even-headway timetable such that the Max load demand is satisfied. Having different vehicle sizes available, the choice of vehicle is sometimes ambiguous. Therefore, three main strategies are considered: • Strategy C1: Minimizing the size of the bus by assigning the largest bus size amongst all available buses such that its seat capacity is less than or equal to the average observed (hence expected) passenger load. • Strategy C2: Maximizing the size of bus by assigning the smallest bus size amongst all available buses such that itsseat capacity is greater than or equal to the average observed (hence expected) passenger load. • Strategy C3: Selecting the vehicle, whose size is closest to the average observed demand, per vehicle, at the Maxload point. The methodology developed has been applied to several sets of real data from Auckland, New Zealand. The Pareto frontier of these results exhibits significant improvement over the current set of departures. If the set of all terminals is denoted as T, the sum of D(k) for all k

			in T, is equal to the minimum number of vehicles required to service the set T. This is known as the fleet size formula. Mathematically, for a given fixed schedule S: $D(S) = \sum_{k \in T} D(k) = \sum_{k \in T} \max_{t \in [T_1, T_2]} d(k,t)$ where D(S) is the minimum number of buses to service the set T.
OPTIMAL RESOURCE ALLOCATION FOR PROJECTS[6]	Carbno Collin	Project Management Journal	This research paper mainly considers and designs mathematical models that weigh numerous factors on delivery schedules, in determining the allocation of resources to different projects. This paper helps us in understanding how to formulate a mathematical model that might help us appropriately allocate optimal number of buses to each route. The given paper helps us understand optimal resourcing and Manfred's distribution. It answers many questions about resource allocation, some of these are: How many resources should be assigned to each project?Are these allocations optimal? Could mathematical models provide insight and assistance into the allocation and scheduling of resources to projects? Given the multitude of factors affecting such decisions, one would expect complex formulas to describe the optimal allocation of resources. Contrary to expectations, the formulas describing optimal allocation in this paper are intelligible and practical in use.
REGULATING BUS MANAGEMENT SYSTEM USING CLOUD PLATFORM[7]	Ranjith Ramesh, YokeshEzhilarasu, PrasannaRavichandran, and Soma Prathibha	International Journal of e-Education, e-Business, e-Management and e-Learning	Here, the authors propose a system in which the number of passengers in a bus stop can be calculated and the bus service can be regulated depending on the passenger's arrival. In this paper, the Authors mainly concentrate on the scheme of the proposed system and resource allocation in cloud using Gossip protocol. This paper deals briefly with calculating the total number of passengers in a stop and regulate the bus service accordingly using dynamic resource allocation in cloud. The proposed system has the following objectives: • To find the passenger population in bus stops using message service and online bus pass. • Regulate the bus frequency depending on the passenger population • Allocating the resources properly in cloud by using gossip protocol. • Using GPS system with cloud to find the delay in buses and indicate the passengers through their cell phone By using the algorithm proposed in this paper, the amount of passengers in the bus stop can be calculated and the buses can be regulated.
THE ALLOCATION OF BUSES IN HEAVILY UTILIZED NETWORKS WITH OVERLAPPING ROUTES[8]	Anthony F.HanNigelH.M.Wilson	Transportation Research Part B: Methodological	This paper addresses the problem of allocating a fleet of buses between routes. A formulation of the problem has been developed which recognizes passenger route choice behavior, and seeks to minimize a function of passenger wait time and bus crowding subject to constraints on the number of buses available and the provision of enough capacity on each route to carry all passengers who would select it. An algorithm has been developed based on the decomposition of the problem into base allocation and surplus allocation components. The base allocation identifies a feasible solution using an (approx.) minimum number of buses. The bus allocation procedure developed in this paper has been applied to part of the Cairo bus system in a completely manual procedure, and is proposed to be the central element of a short-range bus service planning process for that city.
DATA CLEANING: OVERVIEW AND EMERGING CHALLENGES[9]	Xu Chu, Ihab F. Ilyas, Sanjay Krishnan, Jiannan Wang	SIGMOD Tutorial	This paper emphasizes on detecting and cleansing the errors in order to provide accurate analytics and reliable decisions. Three main questions that every technique needs to address: • Error Type (What to Detect): Qualitative error detection techniques are classified according to which type of

			errors are captured.Integrity constraints (ICs), a fractional of first order logic, to capture data quality rules that the database should conform to, including functional dependencies (FDs), and denial constraints (DCs) can be used. • Automation (How to Detect): These approaches can be classified according to whether and how humans are involved in the error detection process. Most techniques are fully automatic. • Where to Detect: Errors can happen in all stages of a business intelligence (BI) stack, for example, errors in the source database are often propagated through the data processing pipeline. While most error detection techniques detect errors in the original database, some errors can only be discovered much later in the data processing pipeline, where more semantics and business logic are available. Error Repairing: • Repair Target (What to Repair?): - trusting the declared integrity constraints, and hence, only data can be updated to remove errors - trusting the data completely and allowing the relaxation of the constraints, for example, to address schema evolution and obsolete business rules - exploring the possibility of changing both the data and the constraints • Automation (How to Repair?): • We can classify the current repairing approaches according to whether and how humans are involved. Some techniques are fully automatic, for example, by modifying the database, such that the distance between the original database O and the modified database M is minimized according to some cost function. Other techniques involve humans in the repairing process either to verify the fixes, to suggest fixes, or to train machine learning models to carry out automatic repairing decisions. • Repair Model (Where to Repair?): • Classifications can be based on whether they change the database in-situ, or build a model to describe the possible repairs. Most proposed techniques repair the database in place, thus destructing the original database. For none in-situ repairs, a model can be built to describe the different ways to repair the underlying database.
OPTIMAL ALLOCATION OF VEHICLES TO BUS ROUTES USING AUTOMATICALLY COLLECTED DATA AND SIMULATION MODELLING[10]	Gabriel E.Sánchez-Martínez,Haris N. Koutsopoulos, Nigel H. M. Wilson	Research in Transportation Economics	Monitoring the service quality of high-frequency bus transit is important both to agencies running their own operations and those contracting out, where performance measures can be used to assess contract penalties or bonuses. The availability of automatically collected vehicle movement and demand data enables detecting changes in running times and demand, which may present opportunities to improve service quality and fleet utilization. This research develops a framework to maximize service performance in a set of high-frequency bus routes, given their planned headways and a total fleet size constraint. Using automatically collected data and simulation modelling to evaluate the performance of each route with varying fleet sizes, a greedy

			algorithm adjusts allocation toward optimality. A simplified case study involving morning peak service on nine bus routes in Boston demonstrates the feasibility and potential benefits of the approach. A potential application is automated detection of routes operating with insufficient or excessive resources.
MODIFIED K-MEANS CLUSTERING[11]	RudraPratap Deb Nath, Hyunjo Lee, Nihad Karim Chowdhury, Jae-Woo Chang	14th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems-KES 2010	In this paper, a set of historical data is portioned into a group of meaningful sub-classes (also known as clusters) based on travel time, frequency of travel time and velocity for a specific road segment and time group. With the use of same set of historical travel time estimates, comparison is also made to the forecasting results of other three methods: Successive Moving Average (SMA), Chain Average (CA) and Naïve Bayesian Classification (NBC) method. The results suggest that the travel times for the study periods could be predicted by the proposed method with the minimum Mean Absolute Relative Error (MARE). The importance of travel time information is also indispensable to find the fastest path (i.e. shortest path according to travel time) that connects the origin and destination. Besides, accurate travel time information also helps delivery industries to progress their service quality by delivering on time. Travel time prediction is based on vehicle speed, traffic flow and occupancy which are extremely sensitive to external event like weather condition and traffic incident.the structure of the traffic flow of a specific road net-work fluctuates based on daily, weekly and occasional events. For example, the traffic condition of weekend may differ from that of weekday. Algorithm: 1. Frequency for each travel time is measured by counting the repetition of that travel time in different records. 2. Define Prediction relation that contains three attributes namely Frequency, Travel-time and Velocity. Each tuple of Prediction relation must contain distinct travel time. 3. Find the greatest value from the Frequency attribute (fmax). If two or more tuples contain the greatest value then find the greatest Travel-time for available highest frequencies. A tuple P (xp, yp, zp) is chosen as a centroid of Cluster1, where xp is the maximum frequency, yp is the corresponding maximum travel-time associated with xp and zp is the velocity associated with travel-time yp. 4. Compare each tuple Ti (xi, yi, zi) of relation Prediction with the centroid P (xp, yp, zp) of Cluster1 by using the following formula: COST (P, Ti) = |xp-xi| + |yp-yi| + |zp-zi|Choose tuple Q (xq, yq, zq) as the centroid of Cluster2, where COST(P,Q) is maximum. 5. Build two clusters where the centroid of Cluster1 is tuple P (xp, yp, zp) and that of Cluster2 is tuple Q (xq, yq, zq). 6. Define the cluster memberships of tuples by assigning them to the nearest cluster representative tuple. The cost is given by Eq.1. 7. Re-estimate the cluster centre (we consider arithmetic mean) by assuming the memberships found above are correct. 8. Step 6 and Step 7 are repeated until no change in clusters. 9. After complete preparation of clusters, desired predicted time is calculated separately for each cluster by using the following formula: $\tau i=\sum fi*ti/\sum fi$ Where, τi is the travel time obtained from i-th cluster, N is the total number of tuple in associated cluster, fi is the Frequency of the i-th tuple, and ti is the Travel_time of the i-th tuple. 10. If τ1 and τ2 are desired travel times calculated from Cluster1 and Cluster2 respectively, then the final

			predicted approximate travel time, T for the road segment in the specific time group is the arithmetic mean of τ1 and τ2 i.e. T= (τ1+τ2)/2.
GENETIC ALGORITHM FOR BUS FREQUENCY OPTIMIZATION[12]	Bin Yu, Zhongzhen Yang, Jinbao Yao	Journal of Transportation Engineering	Bus frequency is the key to public transport management. To enhance the level of public transport dispatching management is an important pathway of improving the service quality of urban public transport and attracting bus passengers. The sum of expenses paid by the passengers waiting for the bus, the sum of expenses paid by the passengers on the bus, the sum of expenses paid by passengers transferring public transport and the sum of expenses paid by variable operator as target function to do both interests of passengers and enterprises. And introduce genetic algorithm to analysis of the model. In this paper, a bi-level programming model for the bus frequency design is presented, which determines the optimal bus frequencies aiming to minimize the total travel time of passengers subject to the constraint on the overall fleet size of each company by accounting for the route choice behaviors of the users. The objective of the lower level is to assign transit trips to bus route network based on optimal strategy. In the upper level, bus frequencies of routes are optimized as a result of passenger assignment. An iterative approach, which consists of a genetic algorithm and a label-marking method, is used to solve the bi-level model. Finally, the model and the algorithms are illustrated with two test examples. The results show that the optimization can improve the local service level of one company, and the proper integration of several companies probably improves the efficiency of resources and the service level of the whole transit system.
CLUSTERING AND AGGREGATING CLUES OF TRAJECTORIES FOR MINING TRAJECTORY PATTERNS AND ROUTES[13]	Chih-Chieh Hung, Wen-Chih Peng, Wang-Chien Lee	The VLDB Journal	In this paper, we propose a new trajectorypattern mining framework, namely Clustering and AggregatingClues of Trajectories (CACT), for discoveringtrajectory routes that represent the frequent movement behaviors of a user. In addition to spatial and temporalbiases, we observe that trajectories contain silent durations, i.e., the time durations when no data points are available todescribe the movements of users, which bring many challengingissues to trajectory pattern mining. We claim that amovementbehavior would leave some clues in its varioussampled/observed trajectories. These clues may be extractedfrom spatially and temporally co-located data points fromthe observed trajectories. Based on this observation, we proposeclue-aware trajectory similarity to measure the cluesbetween two trajectories. Accordingly, we further proposethe clue-aware trajectory clustering algorithm to clustersimilar trajectories into groups to capture the movementbehaviors of the user. Finally, we devise the clue-aware trajectoryaggregation algorithm to aggregate trajectories in thesame group to derive the corresponding trajectory pattern androute. We validate our ideas and evaluate the proposed CACTframework by experiments using both synthetic and real datasets.The experimental results show that CACT is more effectivein discovering trajectory patterns than the state-of-the-arttechniques for mining trajectory patterns. Mining trajectory patterns and routes are very challengingdue to inherent noises and the limitations of trajectory acquisitiontechnology. Generally speaking, a trajectory consistsof sequential data points recording the locations and associatedoccurrence time sampled from the movements of auser. Given the logged historic trajectories of the user, wenot only aim to identify the sequential relationships, alsotermed as the movement behavior, among regions wherethe user frequently passes by but also to construct detailedtrajectory routes that represent these movement behaviors.In this paper, we study the movementbehavior in terms of spatial regions.The locations and

			occurrence time of data points in two trajectoriesare usually not the same, even if these two trajectoriescapture the same movement behavior of the user.The silent duration denotes a time duration when there is no data point presence due to data loss or samplingstrategies employed in forming a trajectory. The movement paths at silent durations are uncertain. By clustering and aggregating trajectories, our proposedframework infers movement paths at silent durations,which is very challenging due to the unique characteristicsof trajectories (i.e., spatial and temporal bias, temporal shiftsand silent durations).While a considerable amount of research efforts on trajectory patternmining have been reported,they mostly focus on discovering frequent sequences of “hot regions” where the user frequently appears, rather than piecing together the routing paths among hot regions. Even worse,due to the existence of silent durations, they are not able togenerate a sufficient number of hot regions, not to mentioncapturing the trajectory routes among the hot regions in thediscovered trajectory patterns. Furthermore, a user usuallyhas more than one movement behaviors hidden in the loggedtrajectories. As a result, the hot regions identified from thewhole collection of a user’s trajectories may be too generalto precisely represent her movement behaviors. To addressthis issue, an idea is to first cluster the trajectories of a userinto several groups. Each group of similar trajectories is supposedto represent one movement behavior of the user. Thus,the hot regions (as well as the trajectory route) derived fromtrajectories in the same group are more representative of theparticular movement behaviour.
SCHEDULE-BASED TRANSIT ASSIGNMENT MODEL WITH VEHICLE CAPACITY AND SEAT AVAILABILITY[14]	YounesHamdouch, H.W.HobAgachaiSumalee ,Guodong Wang	Transportation Research Part B: Methodological	In this paper, we propose a new schedule-based equilibrium transit assignment model that differentiates the discomfort level experienced by sitting and standing passengers. The notion of seat allocation has not been considered explicitly and analytically in previous schedule-based frameworks. The model assumes that passengers use strategies when traveling from their origin to their destination. When loading a vehicle, standing on-board passengers continuing to the next station have priority to get available seats and waiting passengers are loaded on a First-Come-First-Serve (FCFS) principle. The stimulus of a standing passenger to sit increases with his/her remaining journey length and time already spent on-board. When a vehicle is full, passengers unable to board must wait for the next vehicle to arrive. The equilibrium conditions can be stated as a variational inequality involving a vector-valued function of expected strategy costs. To find a solution, we adopt the method of successive averages (MSA) that generates strategies during each iteration by solving a dynamic program. Numerical results are also reported to show the effects of our model on the travel strategies and departure time choices of passengers. 1. This paper formulates a model of transit assignment with seat allocation and strategic decision behaviour. 2. The proposed model is based on a schedule-based dynamic transit assignment model with strict vehicle capacity constraint. 3. Explicit consideration of seat allocation leads to a more realistic departure time profile of transit passengers. 4. The equilibrium model of the complex transit model with seat-allocation and strategic behaviour can be solved analytically.
DYNAMIC TRAVEL TIME PREDICTION MODELS FOR BUSES USING ONLY GPS DATA[15]	WeiFan, Zegeye Gurmu	International Journal of Transportation Science and Technology	The purpose of this research is to develop and compare dynamic travel time prediction models which can provide accurate prediction of bus travel time in order to give real-time information at a given downstream bus stop using only global positioning system (GPS) data. Historical Average (HA), Kalman Filtering (KF) and Artificial Neural Network (ANN) models are considered and developed in this paper. A case has been studied by making use of the three models. Promising results are obtained from the case

			study, indicating that the models can be used to implement an Advanced Public Transport System. The implementation of this system could assist transit operators in improving the reliability of bus services, thus attracting more travellers to transit vehicles and helping relieve congestion. The performances of the three models were assessed and compared with each other under two criteria: overall prediction accuracy and robustness. Historical Average Models models give the current and future travel time from observed historical bus travel time data of previous journeys by assuming that the current traffic condition is to remain stationary. The performance of these models are weak unless the traffic pattern in the area of interest is relatively stable over time or where congestion is minimal. The regression models require a linear mathematical function to explain a dependent variable with a set of independent variables. Unlike the previous models, these are able to work satisfactorily even if traffic conditions are not stable. . They usually measure the simultaneous impact of various factors, which are independent between one and another, affecting the dependent variable. The regression models were outperformed by other type of models. However, these models have a relative advantage in revealing which independent variables are less or more important for predicting travel times. Artificial Neural Networks (ANN)models are able to deal with complex and noise data and are suitable to find nonlinear relationships between dependent variable and independent variables. They can be used for prediction purpose, without explicitly specifying the (physical) traffic processes.ANNs, inspired by biological neural networks, are constructed with multiple layers of processing units, known as artificial neurons. The neurons contain activation functions which are highly interconnected with one another by synaptic weights. Through learning process, the synaptic weights are adjusted to map the input-output relationship for the analysed system automatically . The ANN model outperformed the other two models in both aspects.It is shown that bus travel time information can be reasonably provided using only arrival and departure time information at stops even in the absence of traffic-stream data. Another interesting model discussed int this paper was Kalman filtering model.Thesemodels could be used to predict the future state of the dependent variable. They have elegant mathematical representations (e.g. linear state-space equation) which can adequately accommodate traffic fluctuations with their time-dependent parameters (e.g. Kalman gain). These models have also been used by many authors in bus travel time prediction. The basic function here is to provide estimates of the current state of the system from previous time steps. They can also serve as the basis for predicting future values or improving estimates of variables at earlier times because of their capacity to filter
AN ALGORITHM TO AUTOMATICALLY GENERATE SCHEDULE FOR SCHOOL LECTURES USING A HEURISTIC APPROACH[16]	Anirudha Nanda, Manisha P. Pai, and AbhijeetGole	International Journal of Machine Learning and Computing	This paper proposes a solution for a School Timetabling problem from a teacher's point of view, their availability for a given time slot with a focus on solving the issue of clash of lectures and subjects. The proposed algorithm is a heuristic approach to the problem which tries to find the best possible solution (may not be optimal). It makes use of 3 data structures to identify different scenarios and resolve conflicts. A random subject sequence is generated to fill up the timetable followed by making a check whether the teacher is available for a subject at a particular slot. If yes, it is finalised in the Output data structure, otherwise, it is moved to the Clash data structure. Subjects in the Clash data structure are allocated if possible in the available slots for a day. If not, these are moved to the Day_Clash data structure which has a higher precedence for the timetable for the next day. The algorithm, however, is based on the assumption that there won't be any repetition of subjects. The heuristic algorithm uses a cost

			function that calculates cost based on number of conflicts generated. The algorithm has scope for extension and application to transportation timetabling.
INSTANCE-BASED PREDICTION OF CONTINUOUS VALUES[17]	Tony Townsend-Weber and Dennis Kibler	AAAI Technical Report	The paper talks about using an ensemble of instance based and model based approach for prediction of continuous values. It talks about using k nearest neighbour approach for value prediction by choosing the best possible k and removing irrelevant attributes They demonstrated that using weighted or unweighted methods using either of the two attribute elimination methods improves performance as compared to a controlled unweighted 3 NN.
REAL-TIME OPTIMAL BUS SCHEDULING FOR A CITY USING A DTR MODEL[18]	MakrandWagale, AjitPratap Singh, Ashoke K Sarkar and SrinivasArkatkar		The paper focuses on various costs that are to be considered when designing a model for optimal bus scheduling. These include but are not restricted to bus operation cost, passenger waiting time cost, passenger riding time cost and in-vehicle congestion cost. Optimal frequency of buses was calculated with respect to different factors like bus stop service time, routes, number of buses available.
REAL-TIME OPTIMIZATION MODEL FOR DYNAMIC SCHEDULING OF TRANSIT OPERATIONS[19]	Liping Fu, Qing Liu, and Paul Calamai		The paper proposes a solution using a local bus service that stops at all stops and another express bus service that can skip stops keeping in mind costs for both bus operators and passengers. It mentions 4 categories of flexible routing and scheduling strategies namely: a) Zone Scheduling – Routes are divided into zones. All inbound services stop at all stations in a zone and skip others. Reduced service frequency and increased waiting time and the possible requirement of cross-zone transfer are the disadvantages b) Short turning – Suitable for origin-destination peak zones which are covered by short-turn trips.It is difficult to determine the turn-back point and the route schedule to balance passenger loads among the trips and to minimize the total fleet size and passenger wait time. c) Deadheading – Vehicles are allowed to run empty through a number of stations at the beginning or the end of their routes to save time and hence reduce the headways at later stations. d) Dynamic Stop Skipping – Vehicles that are behind schedule are allowed to skip low demand stops. Disadvantage is that passengers with either their origin or destination stop being skipped have to wait for at least another headway to get service. The paper formulates a dynamic scheduling model with consideration for both operator and passenger. It uses a skipping control for every alternatebus. This ensures a minimum level of service. The objective function includes components for total waiting time of the passengers who arrive after the departure of previous bus assuming random arrival with an average passenger waiting time equal to half the headway, the total waiting time of those passengers who have been stranded by previous bus. It also includes the total in-vehicle time of passengers summed over all Origin-Destination pairs and the total bus trip time. The authors ran a simulation model and got results for sensitivity of the control schemes to various factors like passenger demand, bus travel time variation and headway.

Output:

The result of this literature survey is to help us get a direction in our efforts by allowing us to learn about different algorithms and approaches and their application to problems that are similar to ours in some respect.

Lacuna:

None of the literature surveyed considered only ticket records as their data set. Most of them have different self-defined attributes that are used for clustering, prediction and scheduling buses. Also, the relative order of importance of those attributes isn't defined explicitly. The papers do not consider the problem of allocating appropriate number of buses and scheduling them on a route as a combination, rather they attempt at resolving either one of them at a time. The literature did not cover the idea of using timestamps to identify timeslots when buses are supposed to travel on their assigned routes.

Need of New System:

At present, there is no system in place that consider just the ticket records i.e., data from previous weeks to dynamically alter bus travel routes and timings. The current implementations focus on growing the service by changing routes completely or modifying the types of buses for better service instead of focusing on using the existing resources in an efficient manner.