SCHEDULE OPTIMIZATION OF FREIGHT VEHICLE FLEET USING DATA ANALYTICS

Nammuni Arachchige Chami Madhushika Keerthisinghe

(168035F)

Degree of Master of Science in Computer Science

Department of Computer Science and Engineering

University of Moratuwa

Sri Lanka

December 2018

SCHEDULE OPTIMIZATION OF FREIGHT VEHICLE FLEET USING DATA ANALYTICS

Nammuni Arachchige Chami Madhushika Keerthisinghe

(168035F)

Thesis submitted in partial fulfillment of the requirements for the degree Master of Science in Computer Science

Department of Computer Science and Engineering

University of Moratuwa

Sri Lanka

December 2018

Declaration, copyright statement and the statement of the supervisor

"I declare that this is my own work and this thesis does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text. Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my thesis, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books)."

Signature: Date:

The above candidate has carried out research for the Master's thesis under our supervision.

Name of the supervisor: Dr. HMN Dilum Bandara

Signature of the supervisor: Date:

Name of the supervisor: Eng. Nishal A Samarasekera

Signature of the supervisor: Date:

Abstract

Schedule Optimization of Freight Vehicle Fleet Using Data Analytics

Schedule optimization is a key decision process of fleet management. However, truck and driver scheduling in multi-plant goods distribution is a complex problem due to geographically distributed customer sites and plants, heterogeneity in trucks, driver behavior, varying traffic conditions, and constraints such as working and resting hours for drivers. Moreover, we need to satisfy conflicting objectives such as maximizing order coverage and minimizing of the overall costs. At present context, the scheduling process is typically handled by a fleet manager who is responsible for assigning both the trucks and drivers to meet the confirmed jobs/orders of a given day. Such scheduling usually happens on the evening of the day prior to the order delivery date. As an NP-complete problem, assigning most suitable pair of vehicle and driver while satisfying both company and customer becomes difficult in a situation where there is an increment of total number of orders. We propose an automated, heuristic-based truck and driver scheduling solution which comprises of a rule checker and a scheduler. Rule checker imposes constraints and conditions such as driver and truck availability, delivery time constraints, and operating and resting hours. A scheduler that applies simulated annealing is proposed to cover as many orders as possible while minimizing the overall cost. The utility of the proposed solution is tested using a workload derived from a real-world bulk-cement distribution company. The results show good coverage of orders where the coverage increased by more than 10% compared to manual scheduling while minimizing the total cost by 35%. Furthermore, the solution has flexibility to tolerate exceptions due to breakdowns, traffic congestion, and extreme weather conditions without a considerable impact on most of the already assigned pairs of vehicle and driver to orders.

Keywords: Heavy Goods Distribution, Multi-plant, Simulated Annealing, Truck and Driver Scheduling

Dedication

I am dedicating this thesis to:

My loving parents, Mr. & Mrs. Keerthisinghe who are my pillars of success

My beloved brother, Mr. Chinthaka M. Keerthisinghe who is my admirer

My dearest husband, Mr. Dinesh Madusanke who stands by me with light of hope and support

All the teachers and friend who encourage and support me.

Acknowledgement

First and foremost, I would like to express my sincere gratitude to my research supervisor Dr. H.M.N. Dilum Bandara for the guidance, motivation and patience given through the research period. I wish to thank Eng. Nishal A. Samarasekera for agreeing to serve as my co-supervisor.

I would like to show special gratitude to my evaluation panel members who were more than generous with their expertise and precious time. Thank you Dr. Amal Shehan Perera and Dr. Supunmali Ahangama for agreeing to serve on my evaluation panel.

I extend my gratitude to the Department of Computer Science and Engineering, University of Moratuwa, for giving approval and assistance to conduct the research work. Special thanks go to both academic and non-academic staff of the department for their continued support.

A very special gratitude goes to staff of Nimbus Venture who provided the Bulk Cement Delivery data set by keeping trust on me.

I also gratitude to the University of Moratuwa for the financial support as the research was supported in part by the Senate Research Grant of the University of Moratuwa under award number SRC/LT/2016/14.

Furthermore, I would like to acknowledge the lecturers, friends and family who supported me to make this research success.

Finally, I extend my gratitude to any individual who supported me with a mere word, encouragement, advices and opinions throughout the research project.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ANN	Artificial neural network
ANS	Artificial Neural Systems
BCD	Bulk Cement Delivery
DF	Deficit Function
DSS	Decision Support System
GA	Genetic Algorithm
IBK	Naive Bayes classifier
ILP	Integer Linear Programming
J48	Decision tree
LP	Linear Programming
ML	Machine Learning
NB	K nearest neighbor
PART	Rule based algorithm
RMC	Ready Mix Concrete
RO	Relief Opportunity
SA	Simulated Annealing
SMO	Support vector machine

1. INTRODUCTION

1.1. Motivation

Nowadays schedule optimization plays a major role in the fleet management because it leads to cost reduction and operational efficiency while using limited resources. Scheduling is a decision making process of transportation and distribution sector, as well as in other service industries [1]. Heavy goods delivery industries such as Bulk Cement Distribution (BCD) is characterized by high volume, low value, and perishability [2]. Driver and truck scheduling in such industries need to focus on maximizing both the profitability and order coverage (i.e., number of orders covered on time) to maintain long-term supplier-buyer relationships [3]. Further, decisions related to heavy goods delivery scheduling should ensure the road and driver safety while focusing on their core objectives. While most orders are recurrent, they need to be delivered to geographically dispersed sites with varying demands. Therefore, differences in operating cost and travel time need to be managed too.

The fleet management company needs to allocate an order to the most suitable truck and driver based on a set of considerations such as delivery location, delivery time, vehicle availability, driver availability, regulations related to working hours and resting hours, traffic, and weather conditions. Moreover, truck and driver allocation should focus on increasing customer satisfaction, operational efficiency, balanced driver income, driver safety, and reduction of total cost for the company. Incidents such as changes in delivery time, cancelling the order and confirming last minute orders by customer side, vehicle breakdowns, accidents, unexpected driver availability, and high traffic in the route make the scheduling environment dynamic where the schedules may change or even canceled. Therefore, a fleet management company requires a scalable solution to achieve company's objectives such as maximize the customer satisfaction, operational efficiency, driver satisfaction, and minimizing the total cost. Moreover, such a solution should be capable of covering as many orders as possible while minimizing overall cost and fairly distribute the income among drivers.

The scheduling process is typically manipulated by a scheduling manager who is responsible for assigning both the trucks and drivers to meet the job/order demands. The effectiveness of the scheduling process highly depends on the scheduling manager's prior experiences and the trial and error process, especially for last minute orders. He/she should also track the progress of delivery to keep the process smoother by making relevant adjustments to overcome sudden changes such as delayed/failed trucks and plants, last minute order confirmations or cancellations, change in delivery time, and driver unavailability due to sickness or just no-show. The scheduling manager should assign the most suitable driver and truck for an order based on a set of parameters such as order location, delivery time, driver and truck availability, and other the constraints such as mandatory breaks for drivers and speed limits applicable for heavy vehicles to ensure the safety of drivers.

Heavy goods delivery industry has several notable differences. For example, these jobs tend to be long-distance and multi-day. Hence, it is important to consider the maximum working and resting hours to ensure the safety of drivers. In this case, customers are geographically dispersed and there are recurrent jobs, as well as one time jobs. Having multiple plants makes the problem even more complex, as different scenarios need to be considered to determine the most suitable job schedule that maximizes the coverage and minimizes the total cost. As the number of orders, trucks, and drivers increase the complexity of the scheduling process also increases. Moreover, driver availability and vehicle availability, more flexible delivery times, orders which are not delivered on the same day as loading day, labor laws in the country makes the problem unique. Hence, it is important to automate the process to achieve an optimized solution to sustain and grow the business.

1.2. Research problem

We consider a company which deliver heavy goods from multiple-plants to geographically dispersed delivery sites. In this context, the research question that this project attempted to address is:

Given set of orders O, trucks V and drivers D; how to automatically schedule trucks and drivers to serve as many orders as possible while reducing cost, maximizing customer and driver satisfaction, and efficiency?

There is known confirmed set of orders which is required to deliver in the next day. Currently these orders are scheduled by a scheduling manager with the use of his prior experience of assigning trucks and drivers to orders in the day prior to delivery date of an order. The manual scheduling process becomes complex with the increment of number of orders, drivers and trucks availability and other constraints and conditions. In such a situation, the automation becomes vital. The research focuses on the problem of automating the assigning truck and driver pairs to confirmed orders meeting the company objectives of enhancing order coverage while minimizing the cost and meeting customer expectations.

1.3. Research objectives

To address the above research problem the study focuses on following research objectives:

- Identify prominent parameters and the multiple objectives with constraints for the heavy goods delivery truck and driver scheduling problem
- Construct a suitable solution approach to solve the scheduling problem which assign most compatible truck and driver pairs to orders with the intention of covering as many orders as possible while minimizing the total cost
- Evaluate the performance of constructed solution using a workload derived from real world bulk cement delivery company
- Compare the results with related work on similar case studies

1.4. Outline

The rest of the thesis is organized as follows. Chapter 2 presents the literature review where the research work related to vehicle and driver scheduling and optimization algorithms are discussed. Chapter 3 presents the problem formulation considering parameters of vehicle, driver, order and solution. Further, this chapter discussed about the constraints, conditions and objectives related to the problem. Proposed techniques that consist of a rule checker and scheduler is presented in Chapter 4. Performance analysis based on real-world dataset is presented in Chapter 5, and concluding remarks are presented in Chapter 6.

2. LITERATURE REVIEW

Leung [4] focused on scheduling to achieve specific objectives which aim on allocating scare resources to the activities while optimization of defined key performance indicators of the company. Ronen [5] described the scheduling by considering planning process of the route that certain events are take place within a time frame. Further Pinedo [1] discussed about the scheduling in the context of resource allocation to tasks by considering the given time limits. Moreover, the author highlighted that the manufacturing and service industries consider scheduling as a regular decision making process.

Routing and scheduling problems related to fleet management are complex even in the static environment [6]. Bielli et al. [6] pointed out that combinatorial problems for instance vehicle routing and scheduling under fleet management are notoriously difficult solve even in static environment. Different case studies such as taxi, bus, electric vehicles, train, and ready-mix concrete delivery routing and scheduling are presented in the chapter.

The literature review is constructed by considering different approaches of scheduling. Section 2.1 presents the research works related to vehicle scheduling while defining the vehicle scheduling. The research contribution related to driver/crew scheduling is presented in Section 2.2. Moreover, Section 2.3 presents the combination of vehicle and driver scheduling research aspects. It is important to consider about the optimization of the solution related to driver and vehicle scheduling in order to find a best solution from the pool of solutions. Section 2.4 and 2.5 discuss about two optimization algorithms, Simulated Annealing and Genetic Algorithm, respectively. Finally, Section 2.6 summarizes the discussion points in the literature review.

2.1. Vehicle Scheduling

According to [7], there are four basic components in bus, railway and passenger ferry planning process such as designing the network route, timetabling, vehicle scheduling and driver/crew assignment which is sequential and process simultaneously. The author discussed about how to assign buses to developed timetable beforehand. The objective of scheduling is accommodating passenger demand while covering timetable with the use of minimum number of homogeneous buses. The step function named Deficit Function (DF) approach is used to minimize the number of vehicles in fixed schedules while considering given tolerances for shifting departure times of buses. Further, in [8], author addressed the vehicle scheduling problem by considering trip characteristics and required type of vehicle. The constructed vehicle scheduling algorithm is based on deficit function theory.

Kim et. al [9] discussed about the school bus scheduling where the trips for each school were given. The purpose of the solution approach is to minimize the demanded number of vehicles and the total distance travelled. Constraints considered are flow balanced constraint, travel time limitations, both trip origin and trip destination points are the depot and only eligible bus can assign to the trips. Further the author has customized the objective function and the constraints by considering the homogeneous fleet and heterogeneous fleet.

In [10], authors stated that it is vital to dispatch vehicles in optimal way in the logistics industry With the purpose of minimizing the cost of operation and quality improvement of the service provided. They have considered about assigning a fleet of vehicles from a plant to known customers with a defined demand and soft time windows. Each vehicle in the fleet has the characteristics of fixed storage capacity, an average speed for travel, economical distance travelled per delivery trip and economical daily working period. The objective function of dynamic multi-trip vehicle scheduling problem consists operating time of the vehicles, delivery delays and overdue working time of vehicles. The considered constraints are maximum loading capacity of vehicle, economical travelling time per vehicle, one vehicle per trip constraint, condition of assigning a vehicle if there is a breakdown of assigned vehicle and the condition of vehicle can serve another customer after full-filling the assigned task. In order to solving the problem, authors have introduced two stage solving strategy where tabu search algorithm is used to develop distribution plan at the first stage and research work focused real-time scheduling to make adjustments for h]the plan with the incorporation of local search algorithm.

Vehicle scheduling problem in trucking industry in the context of homogeneous fleet, less than full truck load and single depot is discussed in [11]. The objective of the two target model is to minimize total driving distance and the number of trucks used which will be ultimately reduce the total cost. In this paper, the context of tasks that deliver goods from a depot to destination point which has no time windows is considered. The approach of scheduling vehicles handles constraints of only one car for each task, capacity constraint by using penalty function approach. After that, authors have incorporated natural coding genetic algorithm With the intention of getting an optimal solution while improving computing speed. Moreover, the parameters of genetic algorithm have been customized such as population size as 50, crossover probability as 0.55 and mutation probability 0.006.

In [12], authors demonstrated that usage of Machine Learning (ML) techniques to automate the ready-mix concrete truck dispatching problem. For a given set of customers, depots, vehicles, starting points and ending points, authors have defined an objective function of minimizing the cost incurred. The performance of six ML techniques; Decision Tree, Rules, Artificial Neural Network, Support Vector Machine(SVM), K-Nearest Neighbors (KNN) and NavieBayes (NB) were compared to find a process to get the exact expert decisions for dispatching the RMC trucks. The results show that ML techniques perform well

Davis et al. [13] researched the feasibility of solving truck dispatching problem by developing artificial neural networks. Authors considered about the assigning three delivery trucks from a regional distribution center to six geographically divided delivery zones focusing on trip only. They haven't considered the return trip for the scheduling process. The multi-layer perception net has been constructed by considering the destination, size, weight, handling classification and route.

Hachemi et al. [14] proposed two phases approach for log truck scheduling problem in weekly context which focused on minimizing the total of transportation and unproductive waiting costs. The solution approach consists two phases, where determining destinations of log trucks for seven days from forest areas to wood mills focusing reduction of cost of transportation at the first phase. In the second phase, it considers the minimizing the cost incurred due to unproductive waiting time of daily transportation of logs. Constraint-based local search and constraint programming approaches were tested for scheduling component of the second phase where the constraint programming approach have given slightly better results.

In [15], authors studied the truck scheduling problem in the context of collecting solid waste. The objective of solution approach is to minimize the costs consisting truck operating cost and fixed cost. One trip served only one time, sufficiency of vehicles and conservative flow conditions are considered to acquire a good solution through a heuristic approach which encompass an auction algorithm and a dynamic penalty method

Further, due to traffic congestion, mechanical failures and accidents affect the vehicle schedule which leads to the rescheduling the vehicles [16]. Authors have developed a decision support system (DSS), a systematic procedure for prescriptive decision-making for both single depot scheduling and rescheduling due to disruptions. The key intention of the system is to minimize the involved operation and delay costs, under the condition that finishing uncompleted trips. Optimal solution is obtained from the feasible solutions with the incorporation of combined forward-backward auction algorithm. Amalgamation of crew scheduling for the developed DSS was one of recommendation of authors.

Author of the paper titled on "Perspectives on practical aspects of truck routing and scheduling" mentioned that there are significantly different vehicle routing and scheduling consideration such that passenger transportation (Bus, taxi, Dial-a-ride, school buses), service operations and truck routing and scheduling or cargo transportation [5]. Author classified truck routing and scheduling as listed in Table 2.1.

2.2. Driver/ Crew Scheduling

Li and Kwan [17] defined driver schedule as a plan that has set of shifts (the work carried out by a single driver per day) to cover all the required driver works. Authors focus on public transport driver scheduling problem. They have incorporated fuzzy set theory and genetic algorithm to find the best shift assignment to the schedule while achieving bi objectives of minimizing number of shifts and costs.

Klabjan et al. [18] presented a solution for airline crew scheduling problem concerning minimizing the crew cost focusing assignment of crew itineraries to flights. The focused area is domestic flights of US airlines which has the hub and spoke flight network. The developed algorithm has two stages. The first stage consists of LP relaxation to "quasi" optimality where generating random pairing, solve the LP using primal-dual simplex algorithm and select candidate pairings/columns. The second stage consists of finding an integer solution branch-and-bound algorithm using the selected columns based on the dual information of LP relaxation.

Further, airline crew scheduling problem has been solved using ant-colony optimization based algorithm in [19]. The authors have developed an algorithm with the objective of minimizing collective crew cost which contains duty payment, expenses for resting and cost for under-utilized time while adhere to the constraints of consecutive city constraint, working hour constraint, maximum inactive time for a duty, flight numbers constraint, flying hour constraint for a duty and inactive hour constraint.

In [20], authors considered the scheduling of transport driver in public sector by developing a framework consisting set of shifts assigning to predefined vehicle schedule. With the aim of selecting the best suitable shifts considering efficiency, number of shifts constraint or duration, they have incorporated an Integer Linear Programming (ILP) approach using a specialized branch and bound technique and applied heuristics to adjust shifts by avoiding repetitions and refining costs. Again, they have checked for the genetic algorithm as an alternative for ILP. The incorporation of genetic algorithm is most apparent for larger problem.

Table 2.1: Vehicle scheduling and routing problem classification.

Truck Routing and scheduling	Fleet size	One	
		Multiple	
	Fleet mix	Physical characteristics	Homogeneous
			Heterogeneous
		Cost structure	Identical costs
			Different cost
	Cost component	Routing cost	
		Ownership cost	
		Carrier truck cost	
		Common cost	
		Product sourcing cost	
	Depot	Single	
		Multiple	
	Nature of the demand	Deterministic	Demand size
		Stochastic	Location/ Time
	Operation type	Delivery	
		Pick up	
		Mixed	
	Number of trips		
	Truck route time	Limited	
		Not limited	
	Road network	Directed	
		Undirected	
		Mixed	
	Distance and time	Measured	
		Estimated	
		Mixed	
		Stochastic	
		Cost minimization	
		Distance/ Time reduction	
		Used truck minimization	
		Utility maximization	
		Workload balancing	
		Decrease use of outside carriers	
		Risk minimization	

Chen et al. [21] proposed an approach to solve the large scale crew scheduling problems in public transport with Chinese meal break rule. This approach assigns group of crews (shifts) to vehicle blocks; set of journeys assigned for a vehicle per day with the objectives of shifts minimization and operational costs. They have considered common constraints such as continuous working time, overall working time, overall driving time, meal break time and long break time, as well as special constraint for Chinese meal break rule which enforces a rule of crew's meal break should be taken within predefined time range. Authors considered a Relief Opportunity (RO) which is considered as the time and location of crew breaks and piece of work has been defined as a task between any consecutive ROs to generate potential shifts using heuristic method while considering the aforementioned objectives and constraints.

New multi-objective metaheuristics; tabu search and genetic algorithm were developed considering scheduling problem for bus driver in the public bus transport companies [22]. A population based meta-heuristic optimization algorithm, harmony search has been incorporated to schedule shuttle bus drivers considering hard constraints and soft constraints [23].

Khosravi et al. [24] discussed about the comprehensive approach to handle railway systems' crew scheduling problem. The approach consists three phases where depth-first-search algorithm is applied to create all the viable sequences of trips, pairings in the first phase. In the second phase, optimal pairings are selected using the set covering model by considering the objective of minimizing the total relative cost of execution selected pairings. Assignment of crew groups to optimal pairings is done at phase three by minimizing the total cost which consists of crew cost groups to the pairings and fixed cost of the employment of the crew groups while adhere to the limitations such as covering all the pairings, maximum number of employed crew groups, minimum and maximum period of working shifts and overlapping of assignment for pairings.

Yaghini et al. [25] discussed assignment of train drivers for given train timetable with the intention of maximizing efficiency of the resulting schedule while meeting all constraints. An LP-based neighborhood by combining a tabu search meta-heuristic is used to develop proposed neighborhood structure. The proposed matheuristic algorithm initiates with a feasible solution. A new heuristic method is applied to generate an initial solution.

In [26], author applied the breadth-first search algorithm for solving truck driver scheduling problem. Authors also allocate time windows for driving and resting hours. Further, Goel and Kok [27] proposed a truck scheduling approach considering improving safety aspects of roads and enhancing the working environment of drivers by adhering to European Union regulations for team drivers. Authors considered about scheduling working, driving, breaking and rest periods with the aim of visiting all the locations within the defined time frame using a depth-first-breadth-second search algorithm.

Authors in [28], presented a review of research studies regarding the truck drivers scheduling problem since the year 2000 and to categorize them to four criteria in which the main objective is to offer an overview of truck driver scheduling. They classified the referred articles according to the types of problems (driver, crew, and vehicle), nature of the goods (ordinary, hazardous), work hours regulation types (United States, Europe union, Canada, etc.) and the various constraints required (duty, rest duration, meal break, and safety). According to the classifications, authors stated that most of the researchers concern about the drivers' perspective in terms of work hour regulations, rest and break time, safety aspects and the experience of the drivers.

When it comes to vehicle scheduling in road freight transportation, Goel [26] mentioned that departure and arrival time can be scheduled with some degree of freedom compared to airline crew scheduling, scheduling of rail transport or mass transit systems. Further author addressed the lack of research on restrictions of driver's working hours.

2.3. Vehicle and Driver Scheduling

Prata [29] introduced a solution which supports multiple objectives of covering maximum number of jobs and minimizing the idling time of vehicles between trips while considering the constraints of maximum number of duties, waiting time and working hours. The context of the solution is a single-depot-based integrated vehicle and crew scheduling solution [29].

Huisman et al. [30] discussed about multiple depot vehicle and crew scheduling problem for the given set of trips with fixed planning horizon. The research work is focused on minimizing total of vehicle and crew costs while ensuring the mutually compatible of vehicle and the crew roster. Authors proposed an amalgamated vehicle and crew scheduling approach which gives significant savings compared to following scheduling approach. Two algorithms which are the combination of column generation and Lagrangian relaxation were compared to find a feasible solution where one algorithm has only variables related to crew duties.

Authors in [31] proposed a Simulated Annealing (SA) based approach with a two-phased methodology for simultaneously scheduling vehicles and drivers for a limousine renting company. First, the initial solution is constructed with the incorporation of greedy algorithm with constraint programming techniques. Then, initial solution was improved at the optimization phase which was based on SA. The rental company has a single depot which is the base for all drivers and vehicles. The company covers a $12000km^2$ geographical area of Paris city and its suburbs which partitioned to 26 zones while identifying major demanding destinations such as hotels and airports. These places are approximated to respective zone centers. Further, pre-computed travel times between identified locations are stored in the database. In order to capture type of the day and time range within the day, a threshold value have been set to avoid null value within zone. The time table is developed for confirmed trips in the evening of the day prior to the delivery date. The main objective of the proposed solution is covering as many as possible trip demand while the secondary objectives are to minimizing number of vehicles and drivers and reducing costs by minimizing the number of upgrades. Authors mathematically formulate a set of constraints for capacity, category, features, skills, maximum spread time, feasible sequences, and paring constraints. Due to the complexity of the constraints, a mathematical formulation based on a partial constraint satisfaction problem is used to express constraints. The trips are sorted in decreasing order of duration, subsequently labelled one by one by a driver-vehicle pair that can handle it. After each assignment, a forward checking procedure is applied to prevent future conflicts. The initial schedule is then improved using a SA algorithm, which is able to find the global optimum in a large search space. Within a short time, the solution supplies good quality schedules in which the major part of the trips is assigned. The constraints are all satisfied whereas the operational costs, including the number of resources, the number of upgrades, and the total idle time are reduced. However, the results show that total idle time of drivers have increased in SA-based solution compared to manual scheduling. This research work forms a good basis for our problem though the context is somewhat different.

2.4. Meta-heuristics

According to Osman and Kelly, meta-heuristic provides a framework for an iterative generation process to explore and exploit the search space with the objective of finding efficient near optimal solution [32]. Moreover, meta-heuristics can be defined as the high level strategies that guide search space without getting trapped in confined areas of search space with the goal of finding optimal solution [33]. Simulated Annealing and Genetic algorithms are considered as meta-heuristics which are used in combinatorial optimization problem.

2.4.1. Simulated Annealing

Simulated Annealing (SA) can be describes as a procedure of finding a global optimum by attempting to move from initial solution to one of its neighbors in the given neighborhood structure [34]. It is a naturally inspired meta-heuristic. The concept of SA initiated incorporating the concept of simulation of the annealing of solids and the aim of solving large combinatorial optimization problems [35].

Authors in [34] mentioned that the there is a significant influence from parameters in the SA process in finding solution. The parameters of the SA process are initial temperature, cooling rate, epoch length and stopping condition. Initial temperature is the maximum temperature at the stage of starting annealing process. The value of initial temperature is considerably high [35]. The cooling rate has been introduced to determine the rate of cooling the function from initial temperature in the annealing process to find optimum solution. Epoch length is number of iterations at each temperature level [34]. Stopping condition is the rule that terminate the SA process.

Authors in [34] have explained the typical procedure of SA as shown in Figure 2.2. The SA process starts with the input of initial solution. Then initial temperature and epoch length is defined according to the problem. The neighborhood solutions are generated by slightly changing the initial solution. With the intention of comparing neighborhood solution against initial solution, objective function is used. The objective function can be a minimization or maximization function according to the specific problems. In the Figure 2.2, it is considered minimization problem. If the neighborhood solution is better than the initial solution, it will be the initial solution for next iteration. If the neighborhood solution is not better than initial solution, we keep the initial solution as it is. This process will be repetitive until SA process meets the stopping condition.

In [36], authors examined six algorithms to find an optimum solution in the context of dispatching ready-mix concrete trucks. They revealed that SA outperformed five other algorithms, namely Hill Climbing, K2, Look Ahead Hill Climbing, Repeated Hill Climbing, Tabu Search, and Genetic Algorithm.

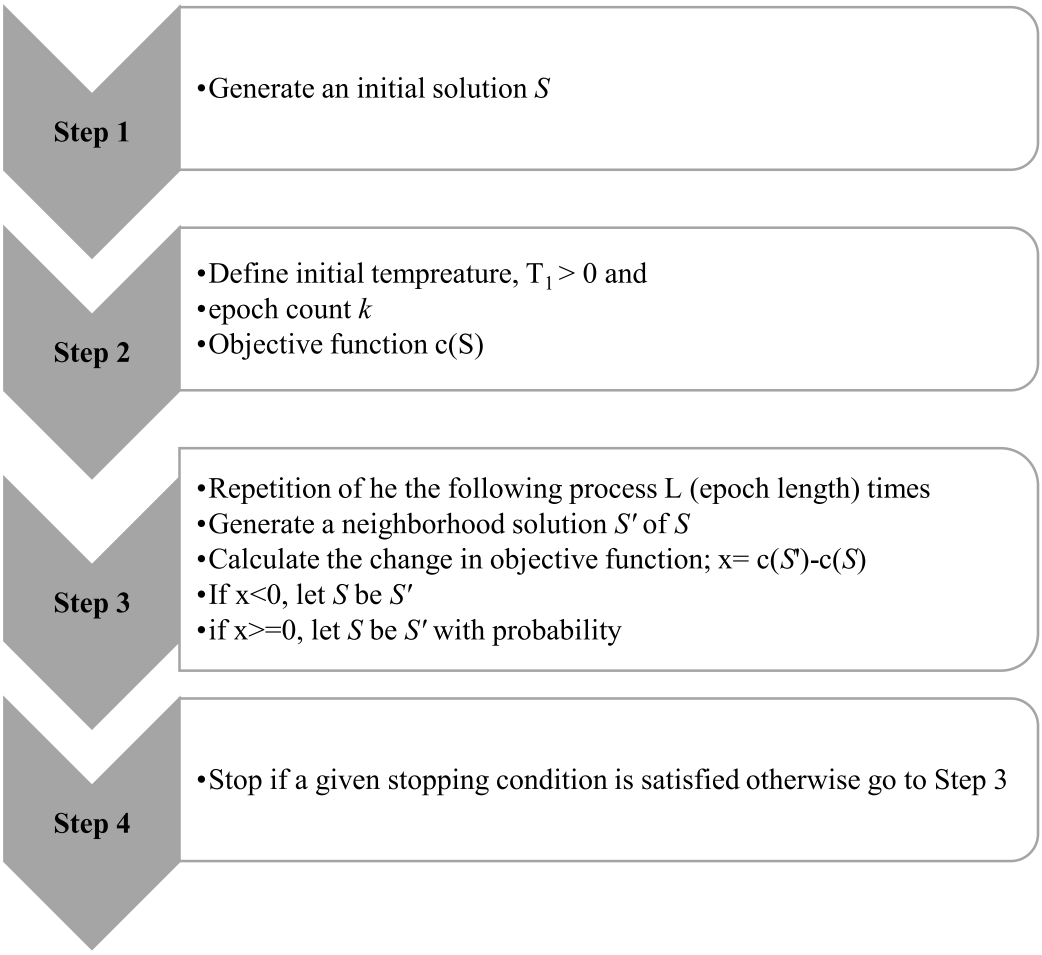

Figure 2.1: Simulated Annealing procedure.

Further SA algorithm has characteristics such as easiness to implement when the neighborhood structure is devised, applicability to wide range of problems, and providing high quality solutions [37]. However, in order to obtain an efficient algorithm, we should carefully define the neighborhood structure and cooling schedule. For a real situation which has special features where there are less likely to have problem specific algorithms, SA can be used with less effort of programming and computing.

2.4.2. Genetic Algorithm

Genetic algorithm is inspired by natural evolution and it is a heuristic search and optimization technique [38]. GA has the base of "Survival of the fittest" concept. Parameters of genetic algorithm includes; the span of individual coding string, population (size of the colony), crossover probability, mutation probability and termination value [39]. Population is the number of feasible solutions that is used in the GA process. An individual solution is considered as a chromosome in GA. Crossover probability is the indicator of how often crossover is performed. The mutation probability decides how often parts of chromosome will be mutated. Authors in [40] explained the process of GA as shown in Figure 2.3. The first step is generating pool of feasible solutions as initial random population for GA process. The population size is determined with the complexity of the problem. When evaluating the fitness for each population, objective function should be defined. After evaluating fitness value, best individuals will be selected for mating pool and conduct the crossover process. Again, fitness value of newly generated solutions will be evaluated and check whether the optimal solution has been found. If the optimal solution is found, the process will stop. Or else the reproduction and crossover process will be conducted and perform mutation.

2.5. Summary

In the context of fleet management researchers have considered different aspects of scheduling. Mainly vehicle and the crew are the main stakeholders of transporting goods from one place to another. Above related work implies that, researches have solved the driver, crew, and vehicle scheduling problems by using various standard and innovative algorithms for many years. There is no ideal way to solve these problems because these problems are NP hard. Moreover, each solution thoroughly depends with the problem context, problem size, and nature of the related constraints.

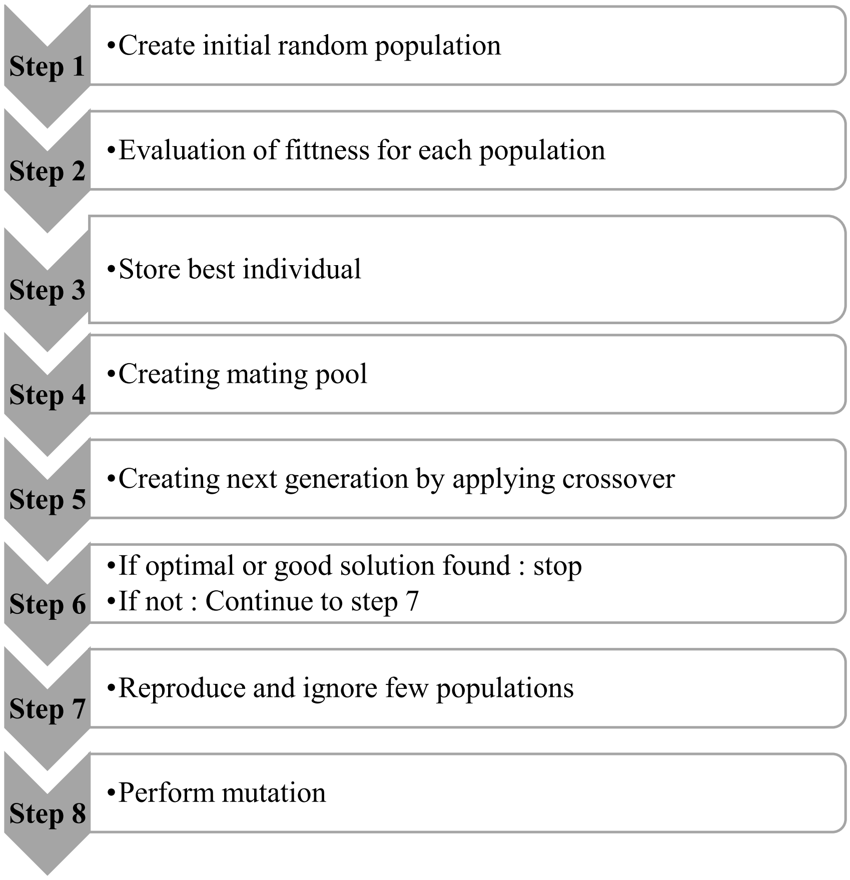

Figure 2.2: Genetic Algorithm procedure.

Several aforementioned related works propose to address the driver and vehicle scheduling problem using different optimization and machine-learning techniques. Therefore, it is crucial to find out an appropriate approach in a given context that provides an acceptable result. Moreover, the chosen approach largely depends on the problem context such as size of the dataset (e.g., drivers and vehicles), number of constraints involved, and accepted output time. The problem we focus on contains a relatively higher number of constraints including multiples set of objectives to be achieved in a priority order. Therefore, it is important to focus on computational time while getting an acceptable solution for truck and driver scheduling problem in the heavy goods delivery context.

3. PROBLEM FORMULATION

The chapter consists of the identified parameters, constraints, and conditions that affect the truck and driver scheduling process and the objective functions. The problem focuses on identifying the most suitable truck and driver to deliver an order while covering as many orders as possible along with minimizing the overall cost. As there are multiple plants, we assume trucks and drivers can be assigned from any plant as per the order demands. Therefore, no explicit home location is assumed for trucks and drivers.

3.1. Characteristics of the problem

Table 3.1 lists the characteristics of the focused scenario of the problem. There are three plants which are situated in geographically dispersed within the country which are capable of catering bulk cement demand. In such case, it is needed to consider which plant is suitable to deliver the pre-determined order. There is more than one vehicle to deliver heavy goods from plant to sites with the same capacity, fuel tank size and model. For a fleet management company, idling vehicles are a cost. Company tries to utilize the fleet by maintaining a drivers pool with higher number than number of vehicles while adhering to working and resting time rules and regulations.

Table 3.1: Problem characteristics.

Description	Characteristic of the Problem
Housing of vehicles/ No of plants	Multiple
Size of Available Fleet	Multiple
Type of Available Fleet	Homogeneous
Capacity of Available Fleet	Same
Total Number of Drivers	Number of Drivers is higher than size of fleet
Nature of demand/ order	Deterministic / 1 Full Truck Load Predefined Delivery Time
Location of demand	Known / at nodes (Geographically dispersed)
Operations	Bulk Cement Deliveries (Drop offs to sites only)
Costs	Travel cost, Driver cost, Vehicle cost

Figure 3.1 summarizes the process of bulk cement delivery which is focused in the study. Let **O** be the set of orders, where each order $o \in$ **O** has a delivery location (o_{DL}) and a delivery time (o_{DT}). These orders are to be processed by a set of trucks **V** and a set of drivers **D**. Each truck $v \in$ **V** has a set of attributes such as fuel mileage with a load ($v_{ave_{mileage}}^{with_{load}}$) and without a load($v_{ave\_mileage}^{no\_load}$). Whereas each driver $d \in$ **D** has a set of attributes such as day and night rates ($d_{hourly\_rate}^{t}$) and maximum working hours per day ($d_{\max\_working\_time}^{day}$) and preferred working days. Moreover, regulatory requirements such as maximum number of driving/working hours a day and minimum resting hours per driver needs to be met.

Figure 3.1: Bulk cement delivery process.

Let us assume the cost per order o is calculated based on the cost of driver c_{driver}, cost of travel c_{travel}, and wear and tear cost $c_{vehicle}$. If an order is not delivered there will be an opportunity cost. c_{travel} of an order depends on the distance driven with and without the load, fuel cost, and the fuel mileage. Whereas the c_{driver} depends on the working hours and hourly rate. $c_{vehicle}$ is assumed to be a percentage of the total distance traveled by truck while the opportunity cost is fixed value that is considered as an average cost per order. Our objective is to cover all orders **O** with trucks **V**, and drivers **D**, such that the cost is minimized across all the orders.

3.2. Problem parameters

Next, we describe each of the parameters and constraints in details, and then formulate the optimization problem.

When customer places an order, it contains an expected delivery date, time and location for each order. Table 3.2 lists the order related symbols. Table 3.3 lists the truck related parameters. Average fuel mileage for truck is different when it is loaded and empty. With that the travel cost will be varied. $v_{ave\_mileage}^{\text{with\_load}}$ and $v_{ave_{mileage}}^{\text{no\_load}}$ have been introduced to capture the different fuel mileages. In order to check the availability of truck to deliver an order, it is important to identify the current status of the truck such as *ON_TRIP*, *IDLE*, and *ON_LEAVE*. The current status of the truck is denoted by the symbolv_{status}^{t}. Table 3.4 lists the driver related parameters. Similar to the truck, we need to check the availability of driver to deliver an order. d_{status}^{t} denotes the current status of driver. Driver's cost is calculated with respect to their working hours. Hence, hourly rate for driver, $d_{hourly\_rate}^{t}$ has been introduced. Table 3.5 lists the solution related parameters. These parameters are described in Section 3.3 and 3.4 further.

Table 3.2: Order related parameters.

Symbol	**Description**
O	Order
o_{DT}	Delivery date and time of order o
o_{DL}	Delivery location for order o

Table 3.3: Truck related parameters.

Symbol	**Description**
V	Truck
$v_{ave\_mileage}^{\text{with\_load}}$	Average fuel mileage of truck v with load
$v_{ave\_mileage}^{\text{no\_load}}$	Average fuel mileage of truck v without load
v_{load_time}	Time taken to load truck v
v_{unload_time}	Time taken to unload truck v
v_{status}^{t}	Truck status at time t

Table 3.4: Driver related parameters.

Symbol	Description
D	Driver
d_{status}^{t}	Driver status at time *t*
$d_{hourly\_rate}^{t}$	Driver hourly rate for given time
$d_{max\_workingl\_time}^{day}$	Maximum working time per day for driver *d*
$d_{cum\_working\_time}^{day}$	Cumulative working hours for the day
$rest(i,j)$	Resting time that a driver is entitled while driving from location *i* to *j*

Table 3.5: Solution related parameters.

Symbol	Description
c_{driver}	Driver payment
c_{travel}	Travel cost of delivering an order
$c_{vehicle}$	Truck's wear and tear cost
$c_{v,d}^{o}$	Cost for delivering order *o* using *v* & *d*
$dist(i,j)$	Distance between location *i* to location *j*
$fuel_{up}$	Unit price of fuel
$o_{Earliest\_Start\_Time}$	Earliest time for starting delivery of order *o*
$o_{Latest\_Start\_Time}$	Latest time for the starting delivery of order *o*
$o_{Earliest\_End\_Time}^{v}$	Earliest time that vehicle v can complete order *o* including trip & return trip
$o_{Latest\_End\_Time}^{v}$	Latest time that vehicle v can complete an order *o* including trip and return trip
$o_{Earliest\_End\_Time}^{d}$	Earliest time that driver d can complete order *o* including trip & return trip
$o_{Latest\_End\_Time}^{d}$	Latest time that driver d can complete order *o* including trip & return trip
o_{time_buffer}	Order delivery time adjustment (+/- hours)
$t(i,j)$	Time taken to travel from location *i* to location *j*
x	Truck wear and tear factor, percentage per unit distance (per km travelled)

3.3. Constraints and Conditions

To be eligible for an order *o*, a driver *d* and a truck *v* need to satisfy the following set of constraints:

Driver and Truck Availability Constraints

A driver has the flexibility of to work on any day. Hence, a driver has three states, namely *ON_TRIP*, *IDLE*, and *ON_LEAVE* according to the availability. To be available for an order, a driver should be either on *IDLE* or *ON_TRIP* state. Therefore, the set of eligible drivers for job *o* can be determined as follows:

$$\forall d \in D; d^{t}_{status} \; != ON\_LEAVE \rightarrow d \tag{3.1}$$

A truck has three states, namely *ON_TRIP*, *IDLE*, and *ON_LEAVE*. A truck that is not available due to maintenance or failure (i.e., status is *ON_LEAVE*) is not eligible to deliver an order. Therefore,

$$\forall v \in V; v^{t}_{status} \; != ON\_LEAVE \rightarrow v \tag{3.2}$$

We calculate the time taken to travel from location i to location j where estimated accurately using services such as the Google Maps API [18] based on plant/delivery location and time, traffic, and weather. Therefore, we define a travel time as *t(i, j)* where the *i* is starting location and *j* is ending location for a trip/return trip.

An order should be delivered within the time frame of $o_{DT} \pm o_{time\_buffer}$, where o_{time_buffer} is the time that an order can be either advanced or delayed from the preferred delivery time, typically with the approval of the customer. Based on this we can derive two starting times to deliver an order, namely $o_{earliest\_start\_time}$ and $o_{latest\_start\_time}$. For example, $o_{earliest\_start\_time}$ can be defined as:

$$O_{earliest\_start\_time} = O_{DT} - t(plant, O_{DL}) - v_{load\_time} \tag{3.3}$$

where *plant* is the place where the goods are loaded and v_{load_time} is the time taken to load the goods to the truck. Similarly, we can derive two job end times; $o^{v}_{earliest\_end\_time}$ and $o^{v}_{latest\_end\_time}$ for a given *v* and $o^{d}_{earliest\_end\_time}$ and $o^{d}_{latest\_end\_time}$ for given driver *d* as follows:

$$o^{v}_{earliest\_end\_time} = o_{earliest\_start\_time} + t(plant, o_{DL}) + v_{load\_time} + rest(plant,\ o_{DL}) + v_{unload\_time} + t(o_{DL}, plant) \tag{3.4}$$

$$o^{d}_{earliest\_end\_time} = o_{Early\_Start\_Time} + t(plant, o_{DL}) + v_{load\_time} + rest(plant,\ o_{DL}) + v_{unload\_time} + t(o_{DL}, plant) + drest(o_{DL}, plant) \tag{3.5}$$

Similarly, $o^{v}_{latest\_end\_time}$ and $o^{d}_{latest\_end\_time}$ can be derived.

Feasibility Constraint for Drivers and Trucks

If the driver *d* is *IDLE* within a given time frame (t_1, t_2), the driver is eligible to deliver an order *o* only if:

$$o_{earliest\_start\_time} < t_1 < o_{latest\_start\_time} \cap o^{d}_{earliest\_end\_time} < t_2 < o^{d}_{latest\_end\_time} \tag{3.6}$$

If the truck *v* is *IDLE* within (t_1, t_2), a vehicle *v* is eligible to deliver an order *o* only if:

$$o_{earliest\_start\_time} < t_1 < o_{latest\_start\_time} \cap o^{v}_{earliest\_end\_time} < t_2 < o^{v}_{latest\_end\_time} \tag{3.7}$$

Working and Resting Hour Constraints

A driver *d* should not exceed the maximum allowed working hours per day, which can be stated as follows:

$$d^{day}_{cum\_working\_time} < d^{day}_{max\_working\_time} \tag{3.8}$$

For regulatory purposes and to ensure the safety and the good health of the driver, it is important to allocate resting time for a driver within the working hours and after completing a trip. It is denoted by $rest(i, j)$.

$$rest(i,j) = \begin{cases} 15mts, \ t(i,j) \leq 270mts \\ (t(i,j) * 0.4872 - 101.544), 270mts < t(i,j) < 1440mts \\ 600mts, \ t(i,j) \geq 1440mts \end{cases} \tag{3.9}$$

3.4. Optimization problem

Given **O**, **V**, and **D**, our primary objective is to cover as many orders as possible. This is required to improve customer satisfaction, as most of the customers in the heave good delivery industry are engaged in long-term business relationships. We want to further satisfy the secondary objective of minimizing the overall cost of deliveries. To minimize the cost, it is important to find the most appropriate (*d*, *v*) pair for an order within the customer requested delivery time frame. Moreover, drivers and trucks should not be idle because it is an unnecessary cost for the company due to not having enough orders. Therefore, the objective function can be formulated as follows:

$$Min \sum_{\substack{o \in O, \\ v \in V}} c_{v,d}^{o} \tag{3.10}$$

where $c_{v,d}^{o}$ is the cost of allocating vehicle *v* with driver *d* for order *o* which can be defined as:

$$c_{v,d}^{o} = c_{travel} + c_{driver} + c_{vehicle} \tag{3.11}$$

$$c_{travel} = \left\{\left(v_{ave\_mileage}^{with\_load} * dist(i,j)\right) + \left(v_{ave\_mileage}^{no\_load} * dist(i,j)\right)\right\} * fuel_{up} \tag{3.12}$$

$$c_{driver} = d_{hourly\_rate}^{t} * d_{travel\_time} \tag{3.13}$$

$$c_{vehicle} = x * c_{travel} \tag{3.14}$$

4. SOLUTION APPROACH

We consider a scenario where deliveries are based on a 24x7 schedule and an order as a full truckload. This does not eliminate the possibility of catering for batch orders as the only requirement is to ensure multiple trucks deliver as per the O_{DL}. We assume that the schedule for next day's delivery orders will be determined on the previous evening based on the confirmed orders, available trucks, and drivers. Here the confirmed order list is fixed and changes for next day order list is not considered at this stage.

As shown in Figure 4.1, the proposed solution consists of a rule checker to enforce the constraints and conditions, as well as an optimization phase that attempts to cover as many as orders possible while minimizing the overall costs incurred. Once the initial solution is generated, SA algorithm is used to optimize the (o, d, v) mapping per order such that the order coverage is maximized and the cost is minimized.

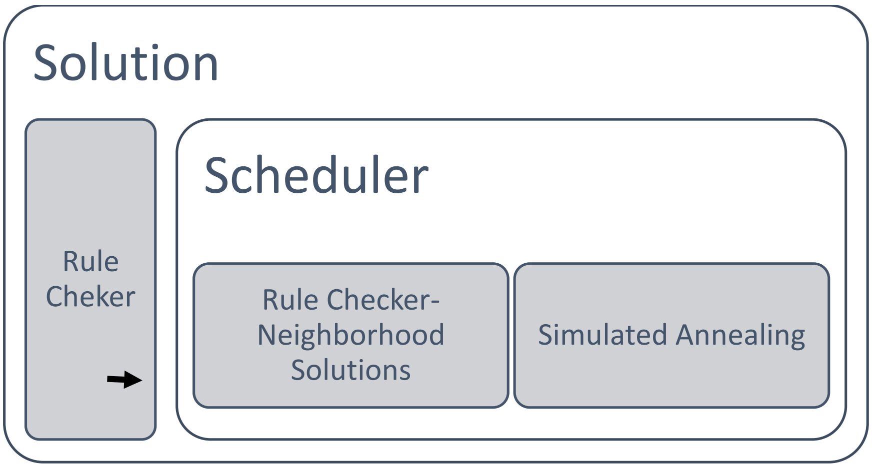

Figure 4.1: Visual representation of solution approach.

4.1. Rule Checker

Given set of orders **O**, vehicles **V**, and drivers **D**, rule checker evaluates the constraints defined in Section 3.3. This reduces the search space as the number of potential trucks and drivers to be assigned to a job depend on availability, delivery time, delivery location, buffer, and resting and operating hours. Because this is an NP-hard problem [41], to achieve an acceptable solution, it is important to reduce the search space. Based on the set of potential vehicles and drivers for a given order, an *initial solution* is derived by randomly assigning a (d, v) pair for an eligible order o.

Figure 4.2 shows the sequential constraint enforcement process and achieving the objectives. The process starts with reducing the search space by checking the availability of driver and the vehicle. Once a driver and a truck are assigned to an order, their status and availability will change. For example, their status will be updated to ON_TRIP for the time range of order delivery including trip and return trip. They cannot be assigned to another order within that period. As mentioned in Section 3.7, drivers and vehicles will be filtered out. Then the earliest start time and the latest start time for delivering order is calculated by considering order delivery time, travel time, resting time for driver and buffer time. Further early end time and late end time is calculated separately for driver and vehicle. Using those parameters, the feasibility of assigned vehicles and drivers will be checked. Finally, it enforces the condition of maximum hours per day for the drivers. Finally, filtered drivers and vehicles are randomly assigned to the selected order. Depending on the availability of drivers and trucks, some orders maybe left unassigned at the end of scheduling. For such orders, there is a fixed opportunity cost for the freight management company which is calculated as an average cost for orders considering the time for delivery as a day and average distance for orders. As the final step of the rule checker, total cost for all orders is calculated. The result from the rule checker is the input for order scheduler.

4.2. Order Scheduler

Order schedule is the second step of the automation problem. After rule checking stage, then the eligible (o, v, d) combinations are further processed through SA algorithm to find an optimal solution while maximizing the order coverage and minimizing the overall cost in the order scheduler phase. In this stage, the initial solution will be changed slightly and go through the rule checker stage to check the eligibility of adjusted assignments of vehicles and drivers to an order. In this solution approach rule checker plays duel role of generating initial rule checker and checking the eligibility of assigned pairs.

SA is used in many global optimization problems in a wide range of domains. The SA algorithm is characterized by an initial temperature, epoch length, cooling function and a cooling rate, and stopping condition [34] .SA algorithm is best known for solving the combinatorial problems [10] and large-scale optimization problems. In such as case, truck and driver scheduling for BCD can be considered as a large-scale optimization problem where it is required to use SA algorithm to find a global optimum solution from a finite number of solutions within reasonable computational time [35]. As mentioned in Section 2.4, the process of order scheduler follows the SA algorithm steps while incorporating rule checker.

Optimization search algorithms mostly optimize a cost or distance matrix. In our solution, our primary objective is to cover all possible orders and then our secondary objective is to optimize cost matrix to minimize the overall order delivery cost. We also able to get a stable solution within acceptable time and results will be discussed in next chapter.

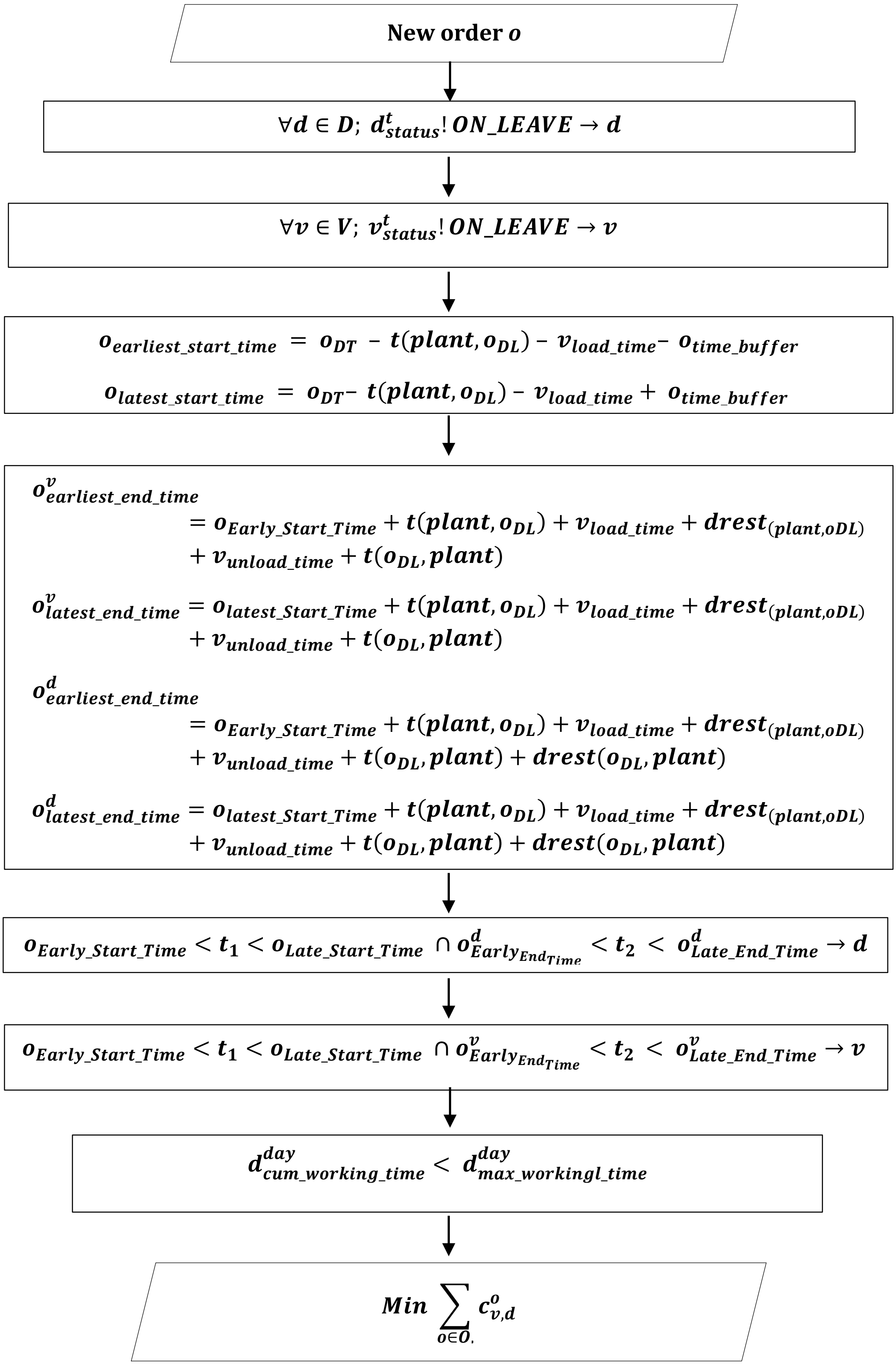

Figure 4.2: Solution process.

5. PERFORMANCE ANALYSIS

The chapter consists of the descriptive analysis of a case: real bulk cement distribution, workload creation for the proposed solution approach incorporating behaviors and patterns of orders, drivers and trucks which were identified from descriptive analysis and detailed discussion of the results obtained from the proposed solution approach. In the Section 5.3, optimized solution is compared with the manual solution with reference to order coverage and cost per km. The impact of buffer hours and delays due to sudden situations to the optimized solution is analyzed. Moreover, the impact to the optimized solution with changes of parameters in SA and GA were discussed.

5.1. Descriptive Analysis of a Case: Real Bulk Cement Distribution

The Table 5.1 shows the characteristics of the dataset received. The selected dataset for fleet consists 13 trucks and one bike, two multiple plants with 53 sites and other 33 service areas as shown in Figure 5.1 for outward and inward journeys between June 1 and 30th June 2016. Out of 13 trucks, six trucks were not used for the analysis due to not having recorded data, error in fuel sensor and vehicle category. The trucks are fitted with GPS devices and fuel sensors. When consider about the fuel sensor type and fuel tank characteristics of the selected trucks, we can conclude that the fleet is homogeneous.

For the descriptive analysis, following parameters were used:

- Timestamp (date and time)
- Longitude
- Latitude
- Ignition status (1 – ignition on or 0 – ignition off)
- Current battery voltage
- Odometer reading

When consider about the data pre-processing, following steps was been taken:

- Latitude and longitude were used to identify the locations such as depots, sites and service areas. The Figure 5.1 shows the identified plants and sites. Further, the data was clustered according to the trips using these points. It can be seen that the customer sites are geographically dispersed.
- In order to calculate fuel level for the given battery voltage in specific time, following formula was used.

$$Fuel\ Level = \frac{(Fuel\ Voltage - Fuel\ Min\ Voltage) * Tank\ Size}{Fuel\ Max\ Volatege - Fuel\ Min\ Voltage} \tag{5.1}$$

- Odometer reading was used to calculate distance traveled between two consecutive points.

Table 5.1: Dataset summary.

	Vehicle id	Fuel sensor type	Fuel max voltage	Fuel min voltage	Tank size (L)	Data Status
1	LY-0234	1-Capacitive	70	6000	500	
2	LY-0354	1-Capacitive	70	6000	500	
3	LY-0547	1-Capacitive	70	6000	500	
4	LY-0549	1-Capacitive	70	6000	500	
5	LY-0552	1-Capacitive	70	6000	500	
6	LY-0560	1-Capacitive	70	6000	500	
7	LY-0561	1-Capacitive	70	6000	500	
8	LY-0551	1-Capacitive	70	6000	500	Error in fuel sensor
9	LY-0558	1-Capacitive	70	6000	500	Error in fuel sensor
10	LY-0121	1-Capacitive	70	6000	500	Data not recorded in csv. File
11	LY-0321	1-Capacitive	70	6000	500	Data not recorded in csv. File
12	LY-0353	1-Capacitive	70	6000	500	Data not recorded in csv. File
13	JN-2075	1-Capacitive	70	6000	500	Not considered as prime mover
14	BBY-6696	0-Analog	0	12000	12	Bike

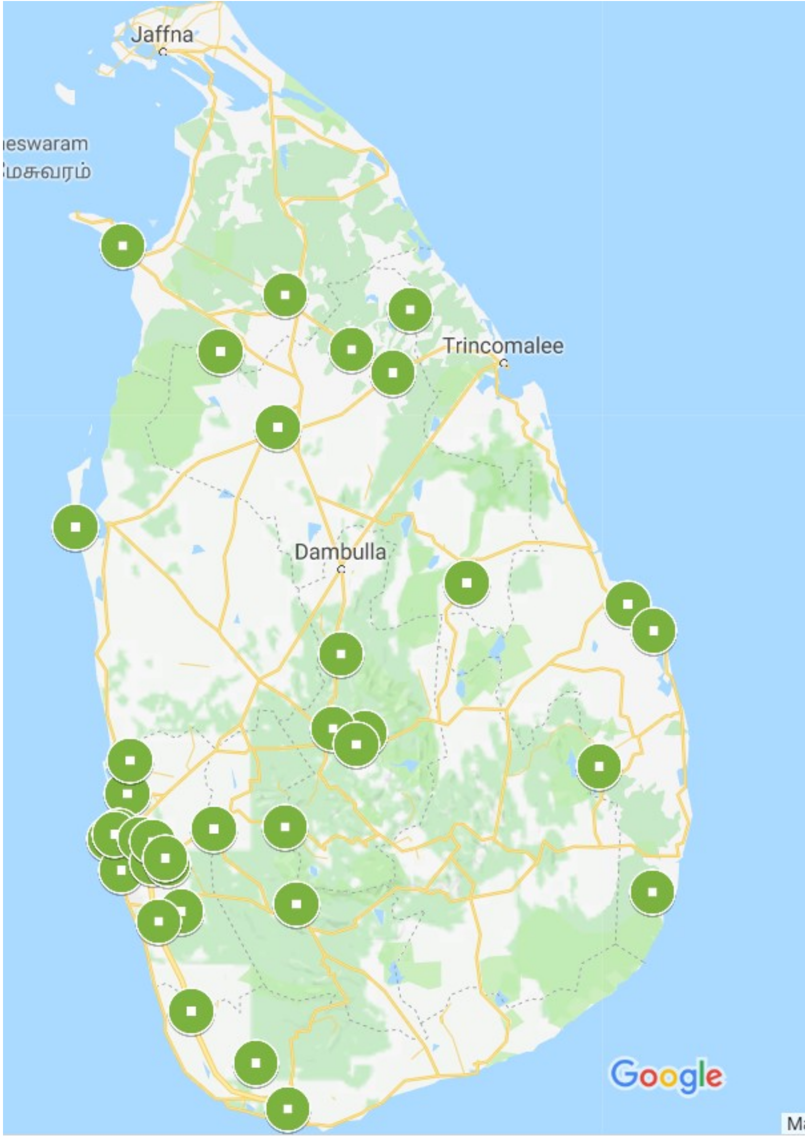

Figure 5.1: Geographical representation of plants and sites.

5.1.1. Trip Summary

Table 5.2 summarize the performance of vehicle in terms of distance, fuel consumption and fuel mileages. Total identified trips for the selected seven trucks was 639 where the fuel mileage deviates between 902.55 Km/L to 0 Km/L. For further analysis, it is decided to consider average fuel mileage of the heavy goods distribution fleet lies between 0.1 Km/L to 6 Km/L. According to the assumption, all the trips which did not comply with the rule considered as outliers. All the trips of truck which has the gray colored record, were filtered out due the fuel mileage is 0.02 Km/L.

Table 5.2: Trip summary with outliers.

With Outliers					
Truck Reg. No	**Total Distance Travelled (Km)**	**Total Fuel Consumption (L)**	**Fuel Mileage (Km/L)**	**Max Fuel Mileage (Km/L)**	**Min Fuel Mileage (Km/L)**
LY-0234	4,676.43	1,699.83	2.75	13.60	0.00
LY-0354	2,762.62	2,982.29	0.93	21.40	0.08
LY-0547	6,238.24	7,796.80	0.80	3.53	0.00
LY-0549	3,669.51	22,121.42	0.17	27.49	0.00
LY-0552	8,643.06	494,180.43	0.02	0.13	0.00
LY-0560	7,106.85	59,430.86	0.12	0.96	0.00
LY-0561	3,806.27	786.07	4.84	902.55	0.25

Total number of trips without outliers was 392 where it represents 61% of the total identified trips. Table 5.3 summarizes the findings after filtered out the outliers. The maximum recorded fuel mileage is 5.53 Km/L while the minimum is 0.10Km/L. Further, Table 5.4 shows the ranges of fuel mileages per truck. The average fuel mileages of all the trucks are less than 3Km/L where the maximum average fuel mileage is 2.91 Km/L.

Table 5.3: Trip summary without outliers.

Without Outliers					
Truck Reg. No	**Total Distance Travelled (Km)**	**Total Fuel Consumption (L)**	**Fuel Mileage (Km/L)**	**Max Fuel Mileage (Km/L)**	**Min Fuel Mileage (Km/L)**
LY-0234	3,962.88	1,633.22	2.43	4.95	0.11
LY-0354	2,711.41	2,862.56	0.95	5.53	0.16
LY-0547	6,232.06	7,697.30	0.81	3.53	0.14
LY-0549	2,658.37	7,923.27	0.34	4.67	0.11
LY-0560	6,293.72	38,227.66	0.16	0.96	0.10
LY-0561	2,012.79	690.72	2.91	5.87	0.25

Table 5.4: Fuel mileage distribution.

Truck Reg. No	Max Fuel Mileage Km/L	Min Fuel Mileage Km/L	Total Fuel Consumption (L)
LY-0234	4.95	0.11	2.43
LY-0354	5.53	0.16	0.95
LY-0547	3.53	0.14	0.81
LY-0549	4.67	0.11	0.34
LY-0560	0.96	0.10	0.16
LY-0561	5.87	0.25	2.91

The filtered trips were analyzed for the aspects such as trip distribution by date, day of the week and day and night. Further vehicle profiles and distance profile were generated to get better understanding about the background of the problem.

5.1.2. Trip Distribution

The average number of trips and return trips per day is 21 where the maximum number of trips were recorded in 29th June 2016 as 42 and lowest number of trips were recorded in 18th and 19th June 2018 as 5. According to the Figure 5.2, we can see a significant reduction of trips and return trips in 18th and 19th June 2016 since those days were mercantile holidays. In 19th June 2016, there were no recorded trips. Compared to the average number of trips there is a 76% of trips reduction in those days. Another observation is there are less no of trips during holidays while there were higher number of trips at the beginning and end of the month.

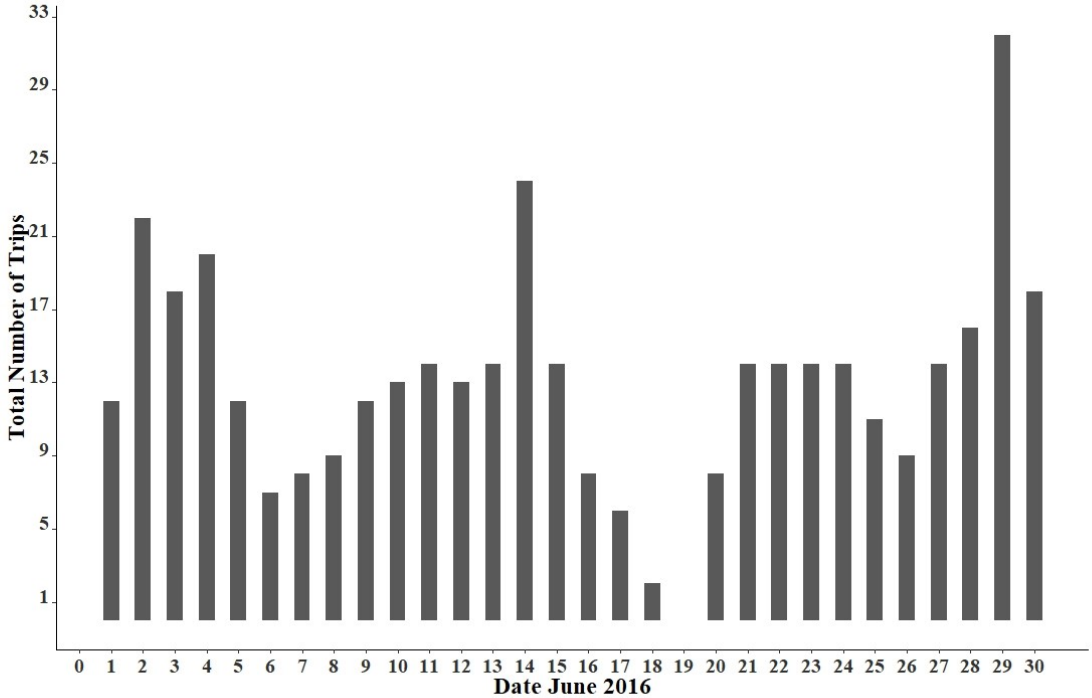

Figure 5.2: Trip distribution by date.

If we consider about the trip distribution which shown in Figure 5.3, according to the day of the week, highest number of trips were recorded in Wednesday while the lowest no of trips was in Sunday. During the mid of the week, trip distribution high while the end of the week, the trip distribution is low. The average trip distribution is shown in Figure 5.4, where Wednesday has highest average number of trips of 20. The average number of trips per week is 98.

Figure 5.5 shows that there is balanced day and night trip distribution in the month where the average number of day or night trips per day is 13. Hence, we can assumed that the BCD industry operates 24x7 schedule.

5.1.3. Distance Profile

When consider about the distance profile as shown in Figure 5.6, the lowest distance was recorded for a holiday while the average distance travelled per day is 823.15 Kms. The recorded lowest distance travelled per date is 44 Kms without considering the 19th June 2016 while the recorded highest distance travelled for a day is 1966.54Kms.

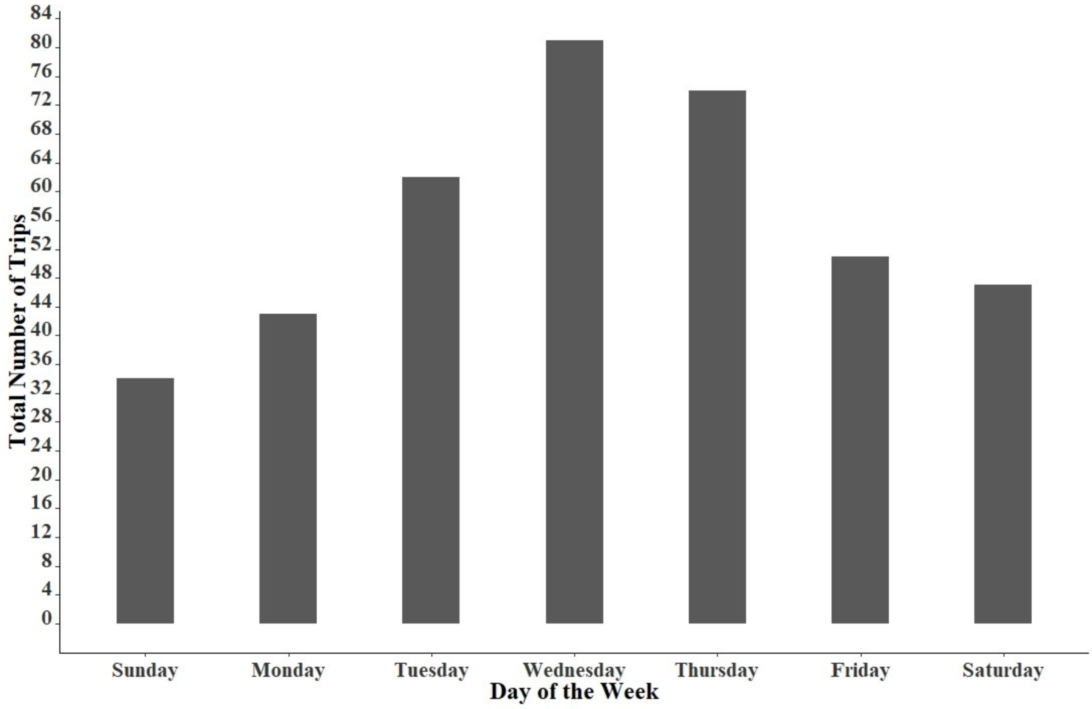

Figure 5.3: Trip distribution by the day of the week.

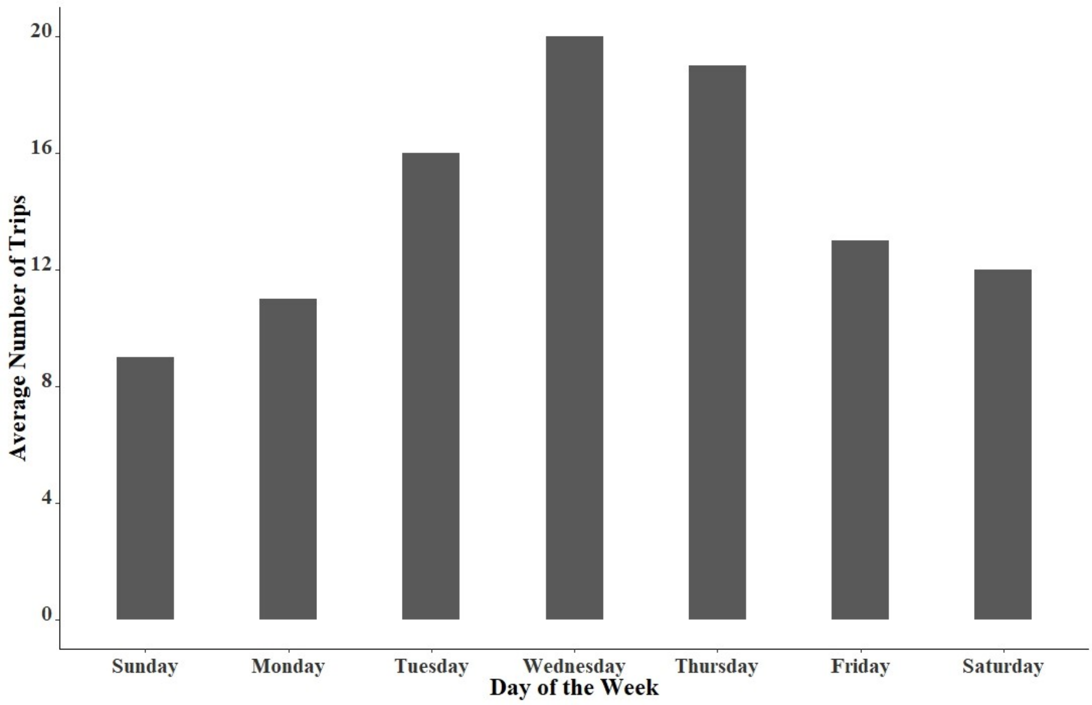

Figure 5.4: Average trip distribution by the day of the week.

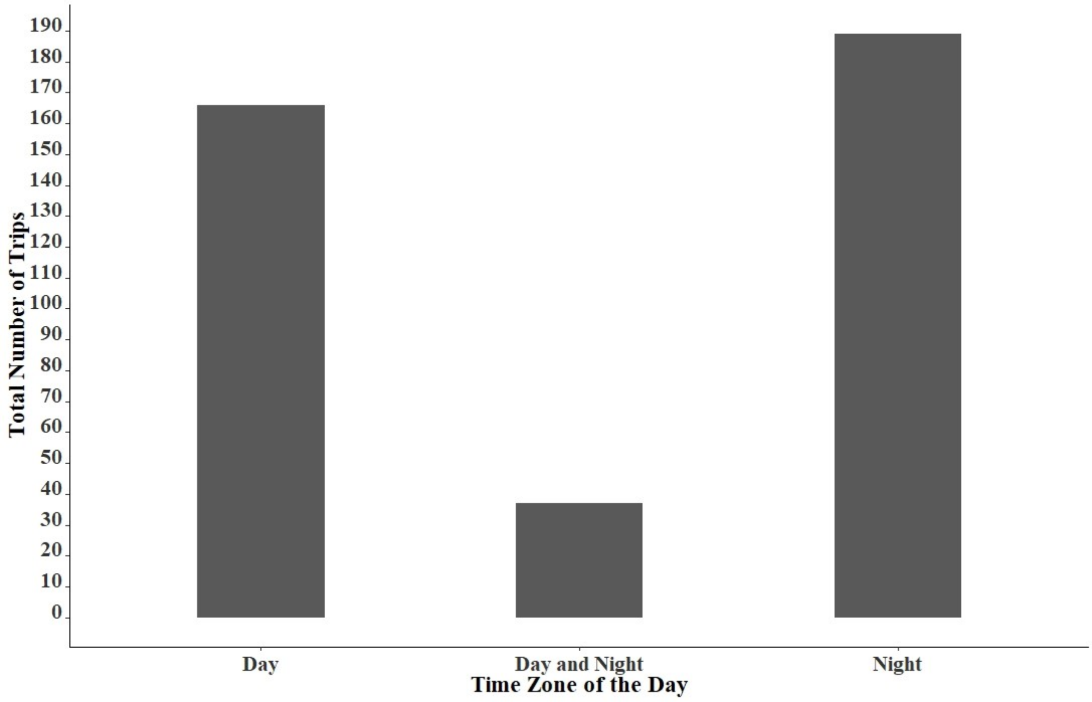

Figure 5.5: Trip distribution by time zone.

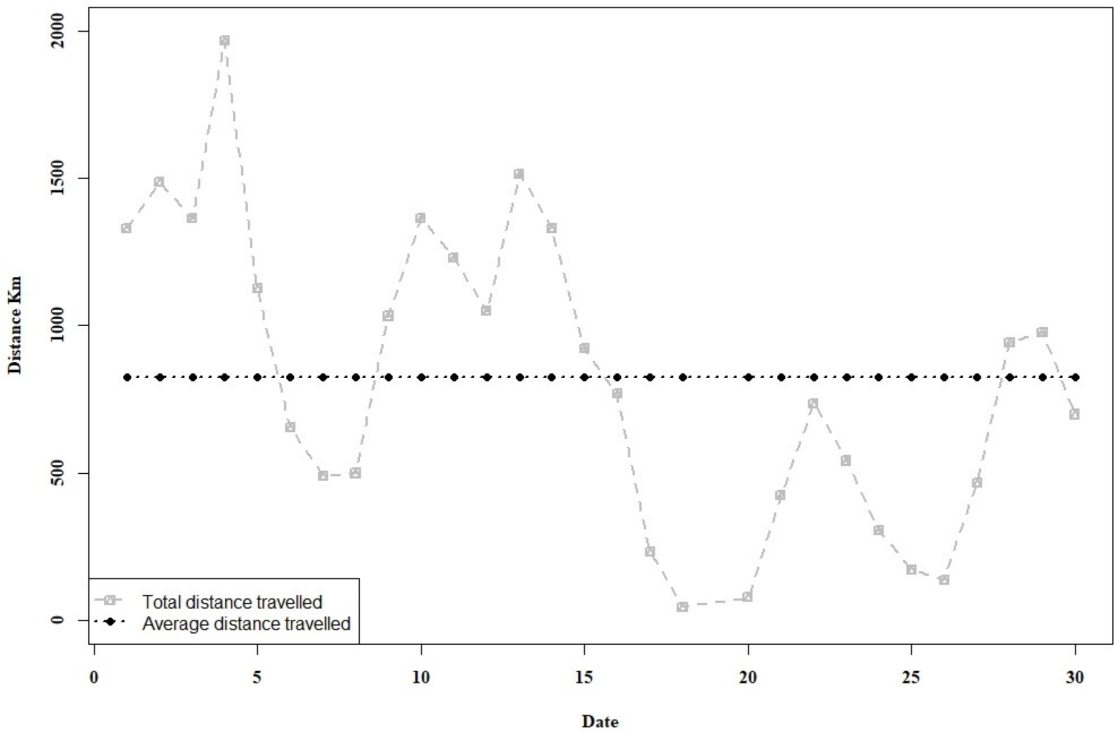

Figure 5.6: Distance travelled by the date of the month of June 2016.

5.1.4. Vehicle Profile

In our solution approach, truck plays a vital role with its performance and the availability. Hence, it is important to identify the vehicle performance in real case. Table 5.5 summarizes the performance of trucks. According to the dataset average trips per day per truck is 3 while there is deviation of operating hours per day. The average operating hours per day per truck is 7-hours. The availability of trucks within a month to deliver orders are different with a range of 19 days to 26 days. The average number of working days per truck per month is 22 days. Figure 5.7 shows the average fuel mileages of respective trucks. The average fuel mileage of a truck lies between 0.0KmL and 3.0KmL. The average fuel mileage per truck is 1.27 Km/L.

In the real scenario, truck and driver is a one entity which Figure 5.8 represents the average operating hours against vehicle. Three of the vehicles are above or equal to 8-hours while three are less.

Table 5.5: Vehicle profile

Truck ID	Distance Travelled (Km)	Fuel Consumption (L)	Average Fuel Mileage (Km/L)	No of Working Days	Operating Hours (H)	Average Operating Hours (H)	No of trips per day
LY-0234	3,962.88	1,633.22	2.43	20	165.45	8	3
LY-0354	2,711.41	2,862.56	0.95	20	132.20	7	4
LY-0547	6,232.06	7,697.30	0.81	24	233.14	10	4
LY-0549	2,658.37	7,923.27	0.34	19	109.54	6	3
LY-0560	6,293.72	38,227.66	0.16	26	236.23	9	3
LY-0561	2,012.79	690.72	2.91	24	103.00	4	2

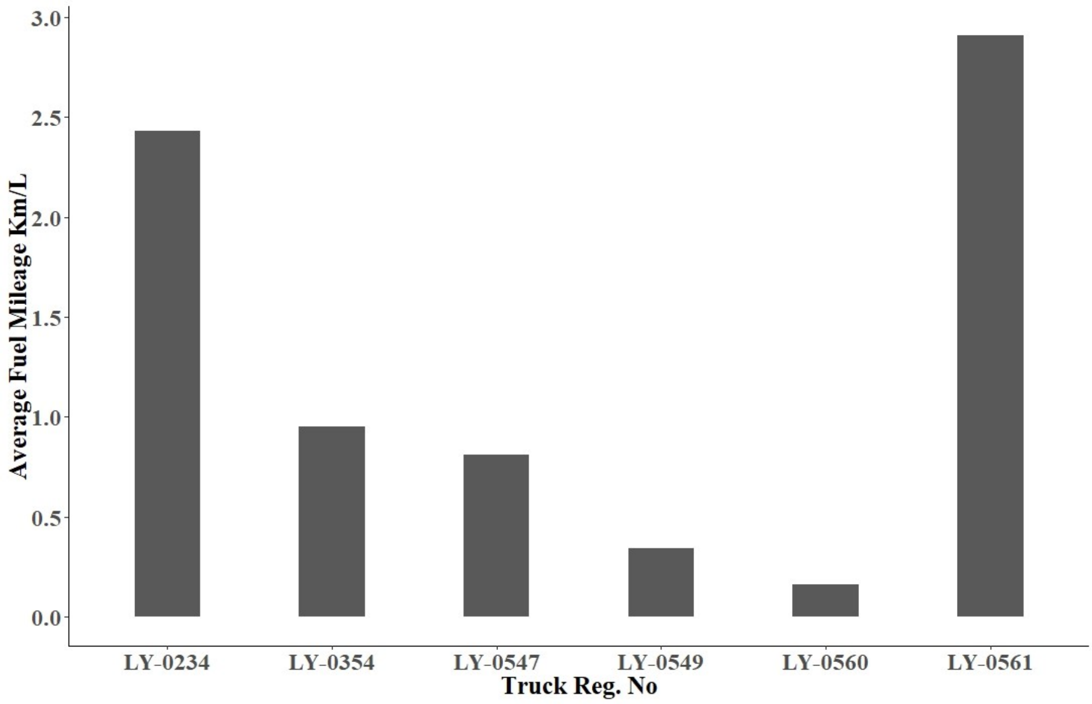

Figure 5.7: Truck vs. fuel mileage.

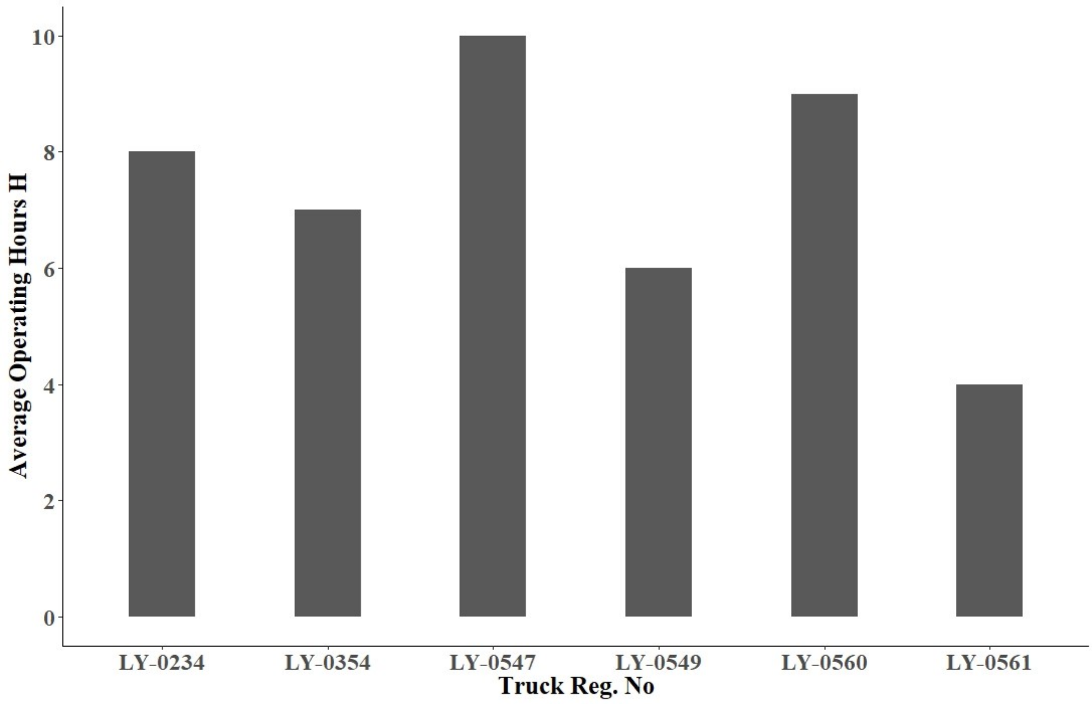

Figure 5.8: Truck vs. average operating hours.

5.2. Workload Creation

A set of synthetic workloads were generated by referring to the behavior of aforementioned real-world BCD Company in Sri Lanka. We considered three plants and multiple sites which are geographically dispersed within Sri Lanka. All distances related to scheduling such as the distance between plant and site, as well as travel times are taken from the Google Distance API [42] to achieve more reliable estimates of distances and travel times.

Table 5.6, 5.7 and 5.8 shows the summary of the synthetic dataset used for the analysis. The total number of trucks in the dataset is 25 and the total number of drivers is 39. Because the plant near Colombo (i.e., commercial capital) has the highest number of order fulfillment, assignment of trucks and the drivers for the plant is matching with the order distribution for the plant. Because the problem runs in a dynamic environment, the results are calculated as a unit cost. We assume that the unit cost for fuel is one, unit daytime hourly rate for the driver as 100, and the nighttime rate is 25% higher than the daytime rate, and the wear and tear factor is set to 10% of the distance traveled.

Table 5.6: Summary of the workload creation.

Day	No of Orders	No of Available Trucks	No of Available Drivers
Monday	38	14	33
Tuesday	52	19	33
Wednesday	70	22	34
Thursday	60	20	32
Friday	46	15	28
Saturday	43	15	27
Sunday	35	13	28

Table 5.7: Truck details.

Truck ID	Fuel Mileage with load Km/L	Fuel Mileage no load Km/L	Load time H	Unloading time H	Mon	Tue	Wed	Thu	Fri	Sat	Sun
V1	2	2.32	0.5	0.5	1	1	1	1	1	1	0
V2	0.7	0.91	0.5	0.5	0	1	1	1	1	0	0
V3	0.6	0.81	0.6	0.6	0	1	0	1	0	0	1
V4	0.53	0.72	0.5	0.5	0	1	1	1	1	1	1
V5	0.5	0.8	0.4	0.4	1	1	1	1	0	0	0
V6	0.35	0.46	0.9	0.9	0	0	1	0	1	0	1
V7	2	2.58	1	1	1	1	1	1	0	1	1
V8	0.75	1	0.5	0.5	1	1	1	1	1	1	1
V9	0.5	0.8	0.5	0.5	0	0	1	1	1	0	0
V10	0.6	0.81	0.6	0.6	1	1	1	1	1	1	0
V11	0.53	0.72	0.5	0.5	0	0	1	1	0	1	1
V12	2	2.75	0.4	0.4	1	1	1	1	1	1	1
V13	0.53	0.86	0.9	0.9	0	0	0	0	1	1	0
V14	0.75	1	0.5	0.5	1	1	1	1	0	1	0
V15	0.5	0.8	0.5	0.5	0	1	1	1	1	1	1
V16	2	2.32	0.5	0.5	1	1	1	1	1	0	0
V17	0.7	0.91	0.5	0.5	0	1	0	1	1	0	1
V18	0.6	0.81	0.6	0.6	1	1	1	0	0	1	1
V19	0.53	0.72	0.5	0.5	1	1	1	1	0	0	1
V20	0.75	1	0.4	0.4	1	1	1	1	1	1	0
V21	0.5	0.8	0.9	0.9	0	0	1	1	1	1	1
V22	2	2.58	1	1	0	1	1	1	1	0	0
V23	0.75	1	0.5	0.5	1	1	1	1	0	1	1
V24	0.5	0.8	0.5	0.5	1	0	1	0	0	0	0
V25	0.6	0.81	0.6	0.6	1	0	1	0	0	1	0

Table 5.8: Driver details.

Driver ID	Available Hours per Day	Mon	Tue	Wed	Thu	Fri	Sat	Sun
D1	8	Idle	Idle	Idle	Idle	Idle	Idle	Idle
D2	7	Idle	Idle	Idle	Idle	Idle	Idle	Idle
D3	8	Idle	Idle	Idle	Idle	Idle	Idle	Idle
D4	12	On_Leave	Idle	Idle	Idle	Idle	Idle	Idle
D5	6	Idle	Idle	Idle	Idle	On_Leave	Idle	On_Leave
D6	8	Idle	On_Leave	On_Leave	Idle	Idle	Idle	Idle
D7	8	Idle	Idle	Idle	Idle	Idle	Idle	Idle
D8	9	Idle	On_Leave	On_Leave	Idle	Idle	Idle	Idle
D9	12	Idle	Idle	Idle	Idle	Idle	Idle	Idle
D10	7	On_Leave	Idle	Idle	On_Leave	On_Leave	Idle	Idle
D11	8	Idle	Idle	Idle	Idle	Idle	Idle	Idle
D12	7	Idle	Idle	Idle	Idle	Idle	Idle	Idle
D13	8	Idle	Idle	Idle	Idle	Idle	Idle	Idle
D14	12	Idle	Idle	Idle	Idle	On_Leave	On_Leave	On_Leave
D15	6	On_Leave	Idle	Idle	On_Leave	On_Leave	Idle	Idle
D16	8	Idle	Idle	Idle	Idle	Idle	Idle	Idle
D17	8	On_Leave	Idle	Idle	On_Leave	On_Leave	Idle	Idle
D18	9	Idle	Idle	On_Leave	Idle	Idle	Idle	Idle
D19	12	Idle	Idle	Idle	Idle	On_Leave	Idle	Idle
D20	7	Idle	Idle	Idle	Idle	Idle	On_Leave	On_Leave
D21	6	Idle	Idle	Idle	Idle	Idle	Idle	Idle
D22	8	Idle	Idle	Idle	Idle	Idle	Idle	Idle
D23	8	Idle	Idle	Idle	Idle	Idle	Idle	Idle
D24	9	Idle	On_Leave	On_Leave	On_Leave	On_Leave	On_Leave	On_Leave
D25	6	Idle	On_Leave	On_Leave	Idle	Idle	On_Leave	On_Leave
D26	8	Idle	On_Leave	Idle	On_Leave	On_Leave	On_Leave	On_Leave
D27	8	Idle	Idle	Idle	On_Leave	Idle	On_Leave	On_Leave
D28	9	Idle	Idle	Idle	Idle	On_Leave	On_Leave	Idle
D29	7	Idle	Idle	Idle	Idle	Idle	On_Leave	On_Leave
D30	8	Idle	Idle	Idle	Idle	Idle	Idle	Idle
D31	7	Idle	Idle	Idle	Idle	Idle	Idle	Idle
D32	8	Idle	Idle	Idle	Idle	Idle	On_Leave	On_Leave
D33	12	Idle	Idle	Idle	Idle	Idle	Idle	Idle
D34	6	On_Leave	Idle	Idle	Idle	Idle	Idle	On_Leave
D35	8	Idle	Idle	Idle	Idle	Idle	Idle	Idle
D36	8	On_Leave	On_Leave	Idle	On_Leave	On_Leave	On_Leave	On_Leave
D37	9	Idle	Idle	Idle	Idle	On_Leave	On_Leave	Idle
D38	12	Idle	Idle	Idle	Idle	Idle	Idle	Idle
D39	7	Idle	Idle	Idle	Idle	Idle	On_Leave	Idle

5.3. Results

The effectiveness of the schedule derived using SA algorithm depends on the cooling rate. Therefore, different combinations were tested on datasets resulting with an initial temperature of 10^4, different cooling rates, and terminating condition of temperature > 1. We ran the simulation six times for each parameter combination while varying the random seed.

Rule checker reduces the search space for the drivers and trucks by enforcing the constraints and conditions. From that search space, a driver and truck will be assigned to order randomly. This assignment for the order list is considered as the *initial solution*. Then SA algorithm is used to find the optimized solution which attempts to maximize the order coverage and minimize the total cost and referred to as the *SA-based solution*.

Since cooling rate increases, the number of instances that scheduler checks for neighborhood solutions increases; there are more chance of finding a better-optimized solution. As seen in Table 5.9, Figure 5.9 and Figure 5.10, as the cooling rate increases the cost is reduced by 50% while the order coverage has increased by 40%. Moreover, the proposed solution based on SA has 40% more order coverage and 60% less cost per unit distance compared to the initial solution. The best order coverage and the minimum cost can be seen when the cooling rate is 0.9. Hence, for further analysis we set the cooling rate to be 0.9.

Table 5.9: Impact of cooling rate.

Cooling Rate		**0.003**	**0.009**	**0.03**	**0.09**	**0.3**	**0.9**
Execution Time (s)		2.4	2.4	3.8	7.1	14.2	135
SA-based Solution	**Order Coverage (%)**	75.7	78.6	78.6	82.9	88.6	88.6
	Cost per km	12.8	10.6	11.4	10.2	8.3	8.3
Initial Solution	**Order Coverage (%)**	28.6	62.9	28.6	47.1	55.7	55.7
	Cost per km	26.4	16.2	27.6	30.4	23.1	23.1

Delivery times are not strict in the heavy goods delivery industry, and customers are willing to tolerate some level of early or late delivery as far as changes are communicated in advance and agreed with the customer. This enables greater flexibility in scheduling jobs such that both the order coverage and profit can be maximized. Therefore, we applied a time buffer by either advancing or delaying the delivery of an order by a predefined time window, as far as it results in a more optimized schedule. To find the impact of varying time windows on scheduling effectiveness, we varied the buffer time as in Figure 5.11 and 5.12. In this case, we tested the datasets throughout the week without changing other parameters. As seen in Table 5.10, order coverage is highest for the buffer time of ± 3-hours. It varied the range of 82% to 100% while resulting in a significant cost reduction of up to 50%.

Because the dataset was created according to the behavior of a real BCD company, we implied patterns of driver availability, truck availability, and the distribution of orders throughout the week. Hence, the demand for orders, availability of trucks and drivers were highest on Wednesday while it was lowest on Sunday. Thursday had the second highest demand. The order coverage and the cost per km depending on the location of the site, distance, and the travel time between the plant and site. If we compare Wednesday and Thursday, driver and truck availability is higher than rest of the days. When analyzing Figure 5.13 and 5.14, we can see that the SA-based scheduler increased the order overage of the two days by 11.4% and 43.3%, respectively while reducing cost by 51.55% and 67.69% compared to the initial solution. As shown in Table 5.10, the lowest and the second lowest number of orders are on Sunday and Monday, respectively. Even though the availability of trucks is less, availability of drivers is higher in those days. Hence, the solution has achieved 94% and 100% order coverage on Sunday and Monday, respectively. When comparing the cost per km for the initial solution, 60% to 80% cost reduction can be seen with the SA-based solution. These results confirm the effectiveness of the proposed SA-based approach to maximize the order coverage while minimizing the cost.

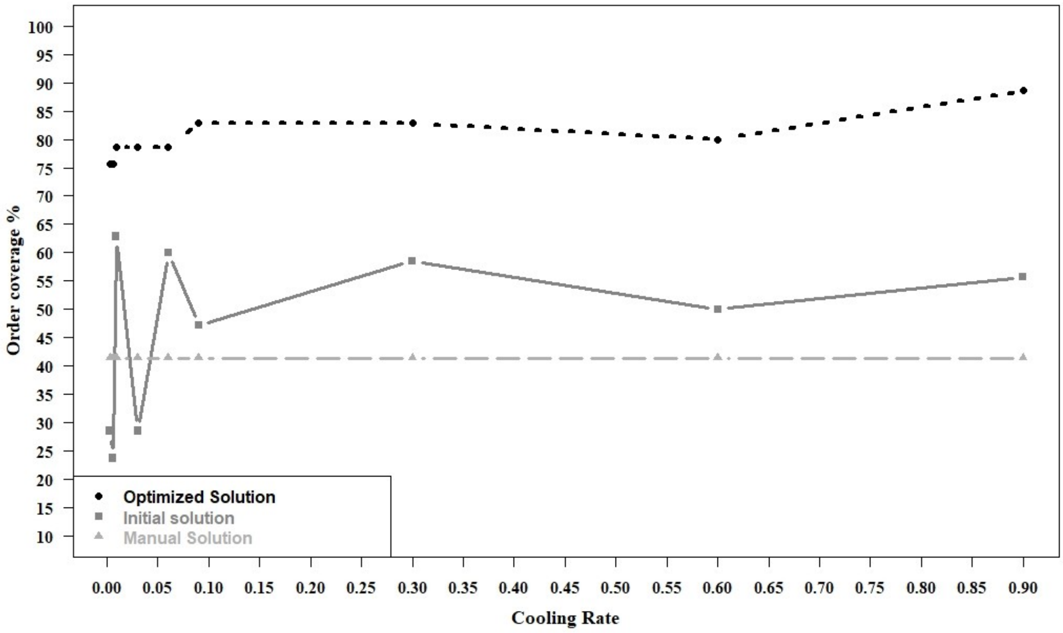

Figure 5.9: Impact of cooling rate in order coverage.

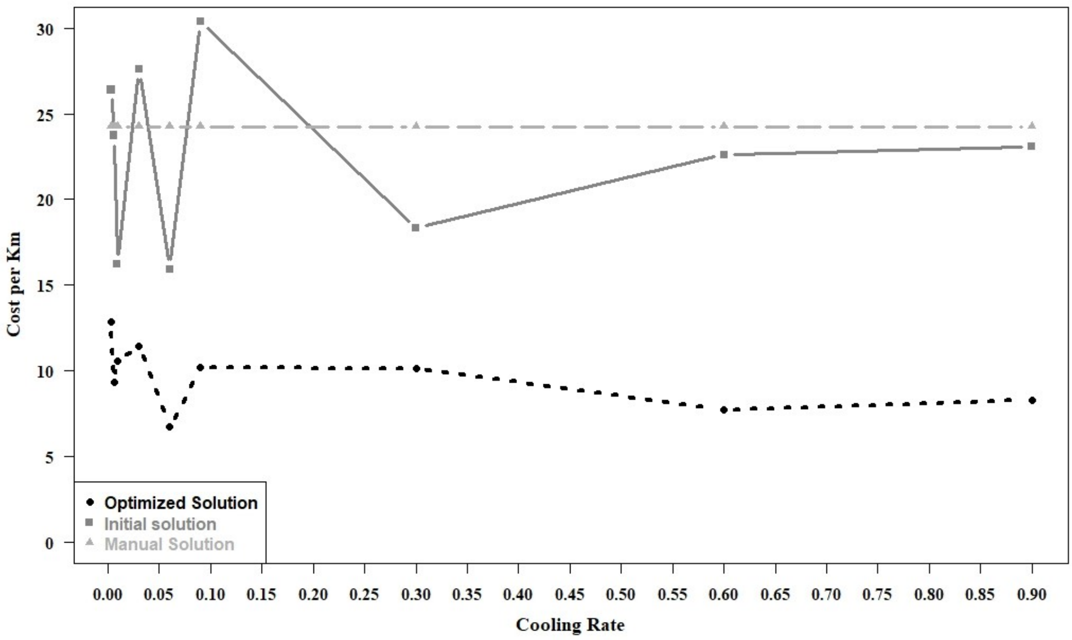

Figure 5.10: Impact of cooling rate in cost per km.

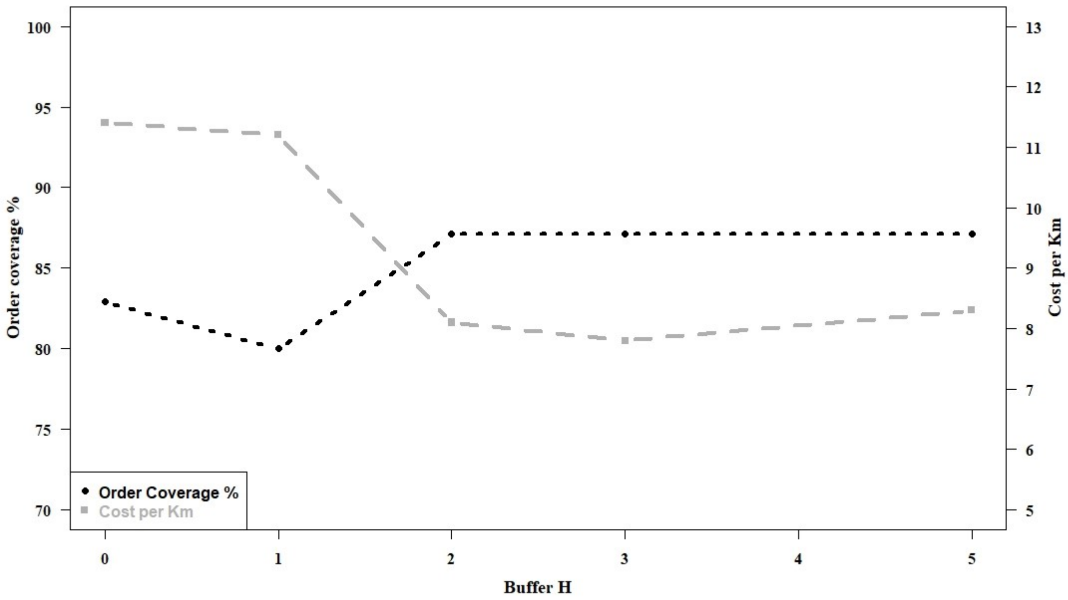

Figure 5.11: Order coverage and cost per km for Wednesday orders (Cooling rate 0.9).

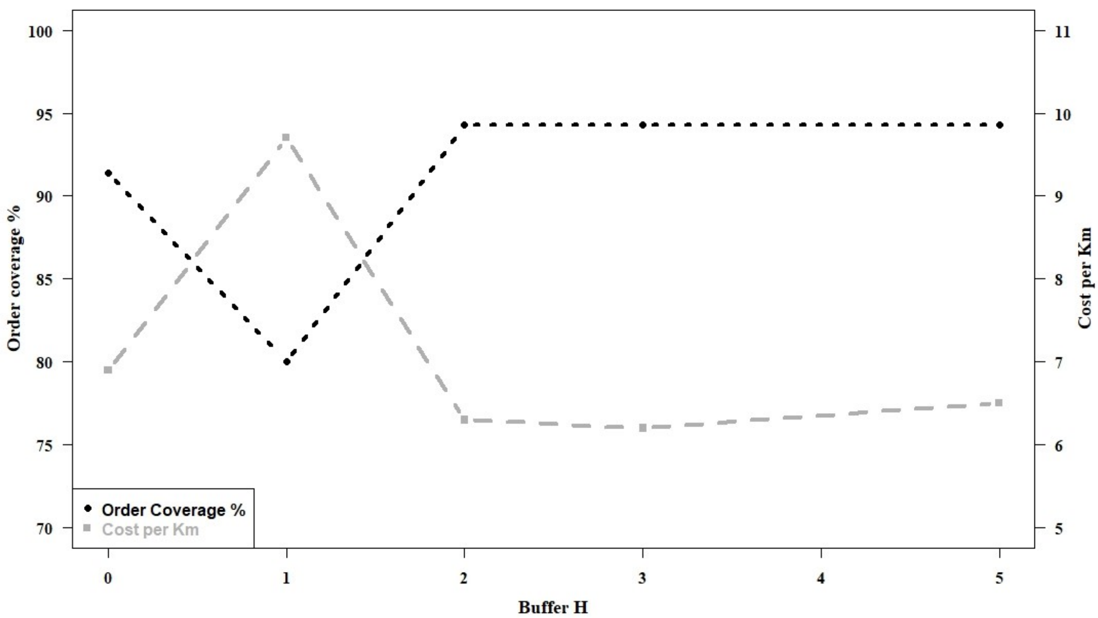

Figure 5.12: Order coverage and cost per km for Sunday orders (Cooling rate 0.9).

Table 5.10: Results for the whole week.

Monday						
Buffer		**0**	**±1**	**±2**	**±3**	**±5**
SA-based Solution	**Order Coverage (%)**	97.4	97.4	100	100	100
	Cost per km	5.7	5.6	5.1	4.9	5.2
Initial Solution	**Order Coverage (%)**	50	50	55.3	47.4	55.3
	Cost per km	24.5	21.5	19.5	25.2	14.4
Tuesday						
SA-based Solution	**Order Coverage (%)**	82.7	90.4	94.2	82.7	90.4
	Cost per km	8.3	6.7	5.5	8.3	6.5
Initial Solution	**Order Coverage (%)**	42.3	55.8	48.1	51.9	57.7
	Cost per km	24.9	18.6	18.2	17.7	15.1
Wednesday						
SA-based Solution	**Order Coverage (%)**	82.9	80	87.1	87.1	87.1
	Cost per km	11.4	11.2	8.1	7.8	8.3
Initial Solution	**Order Coverage (%)**	47.1	41.4	45.7	75.7	67.1
	Cost per km	32.3	30.7	25.1	16.1	25.9
Thursday						
SA-based Solution	**Order Coverage (%)**	75	81.7	75	83.3	76.7
	Cost per km	13.1	9.7	11	9.5	11.9
Initial Solution	**Order Coverage (%)**	41.7	40	50	40	61.7
	Cost per km	29.2	31.1	19.9	29.4	24.8
Friday						
SA-based Solution	**Order Coverage (%)**	80.4	82.6	82.6	93.5	89.1
	Cost per km	9.6	10.3	9.4	6.3	7.9
Initial Solution	**Order Coverage (%)**	36.9	50	52.2	41.3	54.4
	Cost per km	41.6	22.7	27.9	23.8	17.8
Saturday						
SA-based Solution	**Order Coverage (%)**	90.7	86	83.7	86.1	88.4
	Cost per km	6.6	8.2	8.8	7.3	7.8
Initial Solution	**Order Coverage (%)**	48.8	51.2	55.8	58.1	58.1
	Cost per km	25.6	23.1	16.9	16.6	20
Sunday						
SA-based Solution	**Order Coverage (%)**	91.4	80	94.3	94.3	94.3
	Cost per km	6.9	9.7	6.3	6.2	6.5
Initial Solution	**Order Coverage (%)**	51.4	60	57.1	62.7	57.1
	Cost per km	18.2	16.5	15.8	15.8	14.6

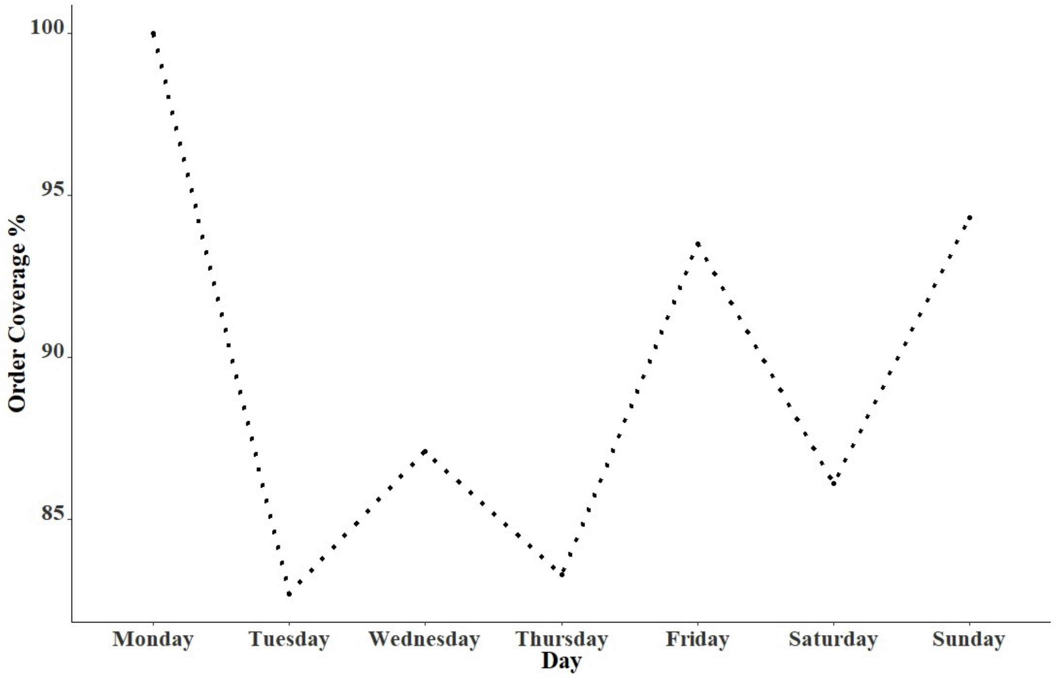

Figure 5.13: Order coverage throughout the week.

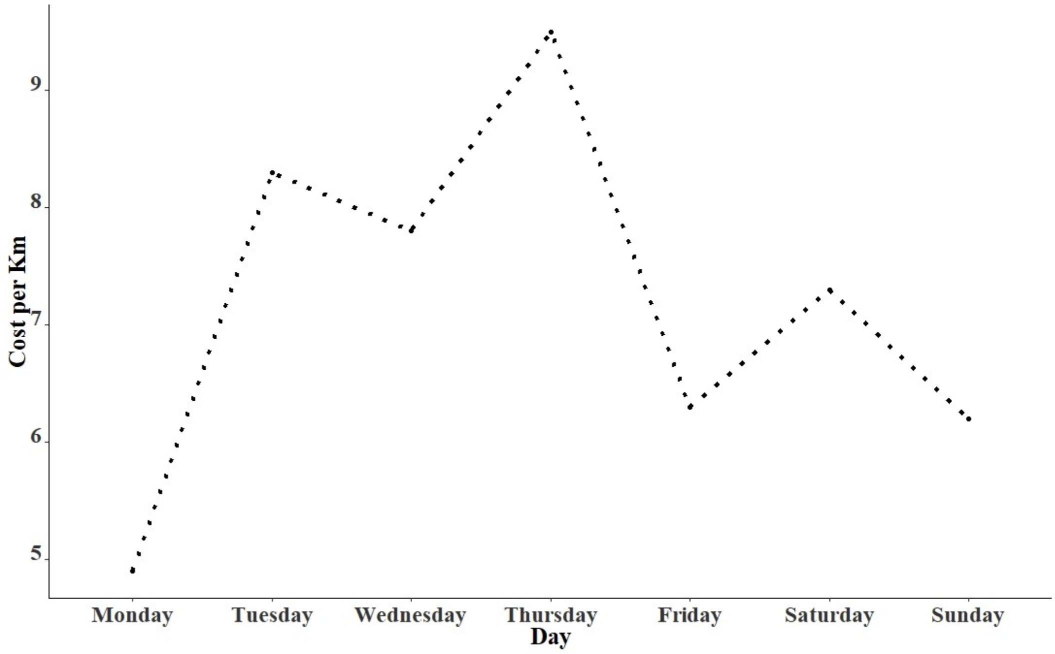

Figure 5.14: Cost per km throughout the week.

Depending on the availability of trucks and drivers, they can be assigned to more than one order within the day. However, due to sudden changes such as traffic, weather, breakdowns, and accidents, all the subsequent orders assigned to such a delayed driver or truck can be affected. To analyze the impact of delayed jobs, 5% of jobs were randomly delayed from 30-minutes to 10-hours for Wednesday’s dataset. Table 5.11 shows the impact due to delayed order completion when the buffer time is ± 3-hours. Because the orders do not have a tight schedule, only a few subsequent jobs get affected. Further, Figure 5.15 shows the graphical representation of the impact of delay. As shown in Figure 5.15 and 5.16, for a delay less than 2-hours the percentage of affected orders are less than 2% while for a delay of 4-hours or more, the percentage of affected jobs is between 2.5% to 7%. This confirms the use of a buffer time (i.e., a time window) not only enhances the order coverage and cost but also enables better tolerance to unexpected delays.

Table 5.11: Impact to the orders due to random delays.

Delay (Minutes)	30	45	60	120	240	480	600
No of orders affected	0	1	1	1	2	4	5
Percentage %	0	1.42	1.42	1.42	2.85	5.71	7.14

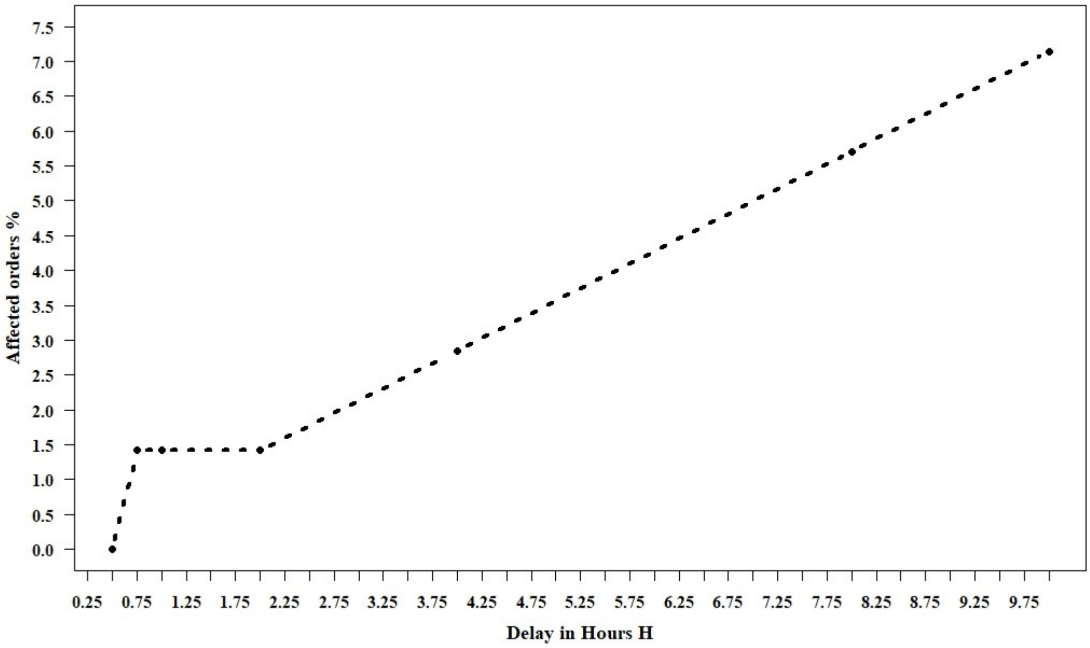

Figure 5.15: Delay vs. number of affected orders.

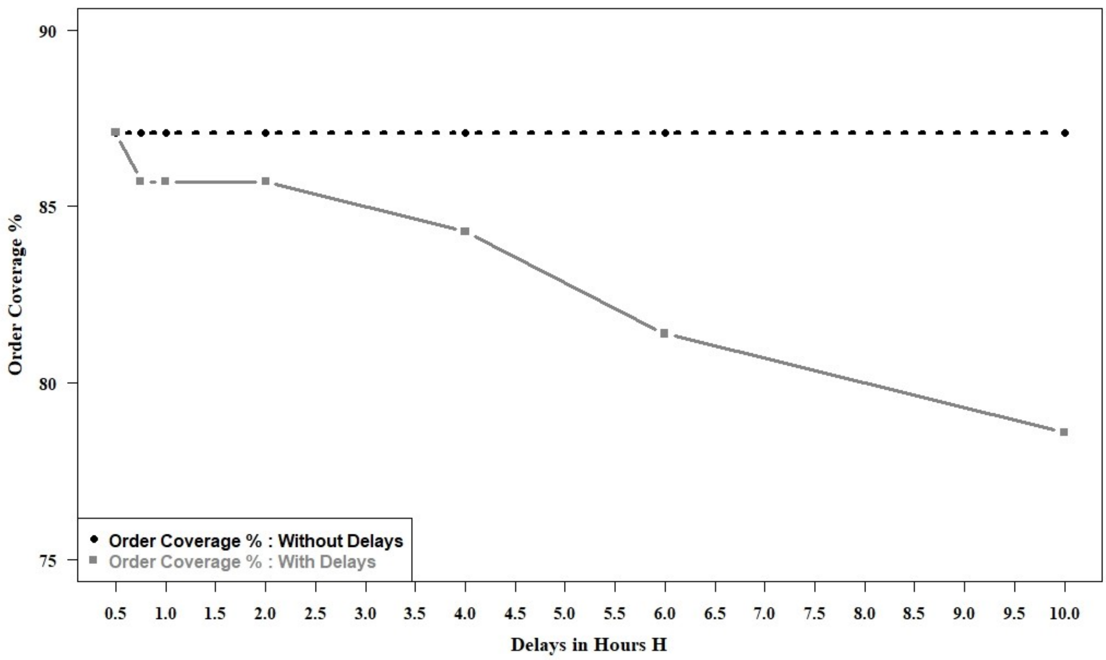

Figure 5.16: Impact of delayed orders against days.

Genetic Algorithm (GA) is another popular technique used to solve scheduling problems; hence, we compare the performance of the proposed SA-based solution with a GA-based solution. GA is a heuristic and stochastic algorithm which uses an iterative process that operates under different parameters such as population size, crossover probability, mutation probability, and the number of iterations [13]. In this research work, GA-based solutions were used for the comparison with proposed solution performance. In the GA-based solution approach, the generated solution after crossovering the chromosomes was checked by rule checker in order to check the feasibility of assigned pairs of vehicle and driver. After that, order coverage and cost per km are calculated to find the optimum solution. The behaviors of population patterns and interdependency of GA parameters such as population size, iterations, crossover probability, and mutation probability were not checked under this research work.

Table 5.12, 5.13, 5.14 and 5.15 show the performance of the GA-based scheduler while using the same initial solution for both SA and GA. When the population size increases, the order coverage increases and cost per km decreases. But the execution time is about one hour. The SA-based solution still has a better order coverage (24.43%) and lower cost (0.5%) compared the GA-based solution. When crossovering two parent solutions, it may result in a set of order allocations which cannot be delivered in the offspring. It is due to there are assigned drivers and trucks which will not satisfy the feasibility constraints of driver and truck in the offspring. Moreover, the execution time of SA-based solution was less than the execution time that was taken by GA-based solutions. Thus, the proposed solution not only has good scheduling properties but also can be executed in modest time enabling even near real-time execution as the orders arrive.

Table 5.12: Impact of population size.

Population Size	Cost Ration LKR per Km	% order Coverage
5	18.70	52.9
10	18.12	52.9
50	17.31	57.1
100	16.11	62.9
500	9.55	68.6

Table 5.13: Impact of crossover probability.

Crossover Probability	Cost per Km	% order Coverage
0	18.41	57.1
0.2	18.12	52.9
0.4	16.53	58.6
0.6	13.18	62.9
0.8	12.40	61.4
1	15.29	58.6

Table 5.14: Impact of mutation probability.

Mutation Probability	**Cost per km**	**% order Coverage**
0	15.86	54.3
0.2	16.10	55.7
0.4	12.64	61.4
0.6	17.28	58.6
0.8	17.02	55.7
1	13.41	60.0

Table 5.15: Impact of number of iterations.

Iterations	**Cost per km**	**% order Coverage**
20	12.64	61.4
50	14.71	60.0
100	10.86	65.7
500	8.14	70.0

6. SUMMARY AND FUTURE WORK

This chapter concludes the study by summarizing the problem formulation, solution approach of schedule optimization of freight fleet using data analytics and the findings of the study. Section 6.3 of the chapter describes the several ways to expand the current study in future work.

6.1. Conclusions

Heavy good distribution truck and driver scheduling in multi-plant goods distribution is a complex problem due to geographically distributed customer sites and plants, heterogeneity in trucks, driver behavior, varying traffic conditions, and constraints such as working and resting hours for drivers. Currently, the scheduling process is a manual process which is handled by a fleet manager who uses his experience. The scheduling task becomes more complex with the increase of number of orders which ultimately leads to time consuming, tedious, more errors and difficulty in finding optimum assignment of vehicles and drivers to orders to achieve company objectives.

In order to cater the aforementioned scenario, we proposed an automated solution. We address the problem of truck and driver scheduling in heavy goods distribution with multiple plants with the objective of maximizing order coverage and minimizing the total cost. The proposed solution consists of a rule checker that enforces various scheduling and regulatory constraints. Rule checker enforces the constraints such as vehicle availability, driver availability, time window and the labor law conditions. The initial solution derived by the rule checker is then optimizes using a scheduler based on Simulated Annealing (SA) algorithm.

Simulation results using a workload trace derived from a real bulk cement distribution company show that the proposed solution assigns trucks and drivers to orders while maximizing order coverage and minimizing cost and computational time. Dataset contains the confirmed order list for a week which align with the real demand pattern for heavy goods distribution. Further it contains the database for drivers and vehicles with their availability for days of week. The solution is capable of covering 10% of orders and reduce cost by 35% compared to manual solution. Moreover, the solution has good tolerance to unexpected delays experienced in the process without causing a major chain reaction. Further the proposed solution was compared with other optimization algorithm, Genetic Algorithm where our solution was outperformed with higher order coverage by 24.3% and less computational time where SA based solution execution time is about maximum 5 minutes while GA based solution execution time is about one hour for scheduling one week confirmed order list. It can be concluded that the proposed solution is capable of covering higher number of orders compared to current manual solution while minimizing the total cost.

6.2. Research Limitations

The proposed solution does not facilitate dynamic orders such as last minute order confirmations, delayed requests, and order cancellations. Because the assignment of trucks and driver to orders are interrelated, the whole schedule may get changed with the dynamic situation. Currently, the solution approach is only consider about the scheduling driver and vehicle when there is a confirmed order list in the day prior to the delivery date.

We have introduced buffer time, flexible time windows for order delivery to overcome the practical situations such as dynamic traffic delays due to congestion, accidents and road maintenance which are. There for we need a mechanism to capture aforementioned real time scenarios without applying buffer time for all the order deliveries.

The dataset does not contain the data related to order capacity where we unable to identify the demand patterns of order capacity. Hence, we consider the homogeneous fleet with the same capacities and delivering full truckload without considering consolidation or break bulk deliveries. The solution approach does not facilitate heterogeneous fleet and multiple deliveries in one trip.

6.3. Future Work

Customer is the main revenue source for a company. It is important to consider about the customer satisfaction with the aim of keeping long term relationships between company and the customers. So, further research can be focused on last minute order confirmation, order cancellation and order delay request to give better customer service. Solution approach can be customized for dynamic order list which enables real time vehicle and driver scheduling in heavy goods distribution industry. By capturing last minute delivery requests arriving within day, we can further improve order coverage while contribution to the company revenue. In order to achieve this requirement, a mechanism to schedule these dynamic orders independently from already scheduled orders.

We consider about the distance and travel time without considering real time scenarios such as delays due to accidents, breakdowns, unexpected traffic and restrictions on road. In such a case, adjustments for unavoidable circumstances can be integrated to the solution approach with a mechanism of capturing the mentioned scenarios.

Our solution focuses on homogeneous fleet and full truckload context. Further research can be conducted considering the heterogeneous fleet and order capacity fluctuations. Our solution approach can be considered as the basic way while introducing the constraint for vehicle category and order capacity. Moreover, a mechanism can be introduced to facilitate the order capacity fluctuations despite of full truckload.

REFERENCES

[1] M. L. Pinedo, Scheduling: Theory, Algorithms, and Systems, 4th ed., New York: Springer Science+Business Media, 2012.

[2] P. Ghemawat, "Distance Still Matters," Harvard Business School Publishing Corporation, 2001.

[3] N. A. C. M. Keerthisinghe, H. M. N. D. Bandara and N. Samarasekera, "Optimization of Truck and Driver Scheduling Using Simulated Annealing," in *IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI)*, Singapore, 2018.

[4] J. Y.T. Leung, Handbook of Scheduling: Algorithms, Models, and Performance Analysis, Chapman and Hall/CRC , 2004.

[5] D. Ronen, "Perspectives on practical aspects of truck routing and scheduling," *European Journal of Operational Research ,* vol. 35, pp. 137-145, 1988.

[6] M. Bielli, A. Bielli and R. Rossi, "Trends in Models and Algorithms for Fleet Management," *Procedia Social and Behavioral Sciences,* vol. 20, pp. 4-18, 2011.

[7] A. Ceder, "Public Ttansport Timetabling and Vehicle Scheduling," in *Advanced Modeling for Transit Operations and Service Planning*, 2002, pp. 30-57.

[8] A. Ceder, "Public-transport vehicle scheduling with multi vehicle type," *Transportation Research Part C,* vol. 19, pp. 485-497, 2011.

[9] B.-I. Kim, S. Kim and J. Park, "A school bus scheduling problem," *European Journal of Operational Research,* vol. 218, pp. 577-585, 2012.

[10] M. Lang, Y. Wang and X. Zhou, "A Two-Stage Algorithm for a Dynamic Multi-Trip Vehicle Scheduling Problem," in *2010 WASE International Conference on Information Engineering*, 2010.

[11] C.-Q.Xu, H. Xuan, B. Li and F.-F. Su, “Vehicle Scheduling Problem for the Single-depot Based on Genetic Algorithms,” in *International Conference of Information Science and Management Engineering*, Xi'an, 7-8 Aug. 2010.

[12] M. Maghrebi, C. Sammut and S. Waller, “Feasibility Study of Automatically Performing the Concrete Delivery Dispatching through Machine Learning Techniques,” *Engineering, Construction and Architectural Management,* vol. 22, no. 5, pp. 573--590 , 2015.

[13] D. L. Davis, E. L. Gillenwater and J. D. Johnson, “An Artificial Neural Systems Framework for Delivery Truck Scheduling,” *IEEE,* pp. 327-333, 1990 .

[14] N. E. Hachemi, M. Gendreau and L.-M. Rousseau, “A heuristic to solve the synchronized log-truck scheduling problem,” *Computers & Operations Research,* vol. 40, pp. 666-673, 2013.

[15] J.-Q. Li, D. Borenstein and P. B. Mirchandani, “Truck scheduling for solidwaste collection in the City of Porto Alegre, Brazil,” *Omega The International Journal of Management Science,* vol. 36, no. 6, p. 1133 – 1149, 2008.

[16] J.-Q. Li, D. Borenstein and P. B. Mirchandani, “A decision support system for the single-depot vehicle rescheduling problem,” *Computers & Operations Research,* vol. 34, p. 1008–1032, 2007.

[17] J. Li and R. S. Kwan, “A fuzzy genetic algorithm for driver scheduling,” *European Journal of Operational Research,* vol. 147, p. 334–344, 2003.

[18] D. Klabjan, E. L. Johnson, G. L. Nemhauser, E. Gelman and S. Ramaswamy, “Solving Large Airline Crew Scheduling Problems: Random Pairing Generation and Strong Branching,” *Computational Optimization and Applications,* vol. 20, no. 1, pp. 73-91, 2001.

[19] G.-F. Deng and W.-T. Lin, "Ant colony optimization-based algorithm for airline crew scheduling problem," *Expert Systems with Applications,* vol. 38, p. 5787–5793, 2011.

[20] A. Wren and D. O. Wren, "A genetic algorithm for public transport driver scheduling," *Computers Ops Res.,* vol. 22, no. 1, pp. 101-110, 1995.

[21] C. Shijun, S. Yindong, S. Xuan and C. Heming, "A Crew Scheduling with Chinese Meal Break Rules," *Journal of Transportation Systems Engineering and Information Technology,* vol. 13, no. 2, pp. 90-95, 2013.

[22] H. R. Lourenço, J. P. Paixão and R. Portugal, "Multiobjective Metaheuristics for the Bus Driver Scheduling Problem," *Transportation Science,* vol. 35, no. 3, pp. 331-343, 2001.

[23] Z. A. Shaffie, Z. A. Abas, A. F. N. A. Rahman, A. S. Shibghatullah and B. Hussin, "Optimization in Driver's Scheduling for University Shuttle Bus using Harmony Search," in *International Symposium on Research in Innovation and Sustainability 2014 (ISoRIS '14)*, Malacca, Malaysia, 15-16 October 2014.

[24] A. Khosravi and M.-N. Tamannaei, "A Comprehensive Approach for Railway Crew Scheduling Problem (Case Study: Iranian Railway Network)," *International Journal of Transportation Engineering,* vol. 4, no. 3, pp. 197-210, 2017.

[25] M. Yaghini, M. Karimi and M. Rahbar, "A set covering approach for multi-depot train driver scheduling," *Journal of Combinatorial Optimization,* vol. 29, no. 3, pp. 636-654, 2015.

[26] A. Goel, "Truck Driver Scheduling in the European Union," *TRANSPORTATION SCIENCE,* vol. 44, no. 4, pp. 429-441, 2010.

27] A. Goel and L. Kok, "Efficient scheduling of team truck drivers in the European Union," *Flex Serv Manuf J ,* vol. 24, p. 81–96, 2012.

[28] M. Koubâa, S. Dhouib, D. Dhouib and A. E. Mhamedi, "Truck Driver Scheduling Problem: Literature Review," *International Federation of Automatic Control,* Vols. 49-12, p. 1950–1955, 2016.

[29] B. d. A. Prata, "A multiobjective metaheuristic approach for the integrated vehicle and crew scheduling," *Journal of Transport Literature,* vol. 10, no. 2, pp. 10-14, 2016.

[30] D. Huisman, R. Freling and A. P. Wagelmans, "Multiple-Depot Integrated Vehicle and Crew Scheduling," *Transportation Science,* vol. 39, pp. 491-502, 2005.

[31] B. Laurent and J.-K. Hao, "Simultaneous vehicle and driver scheduling: case study in a limousine rental company," *Computers & Industrial Engineering ,* vol. 53 , p. 542–558, 2007.

[32] I. H. Osman and J. P. Kelly, "Meta-Heuristics: An Overview," in *Meta-Heuristics Theory and Applications*, Massachusetts, Kluwer Academic Publishers, 1996, pp. 10-30.

[33] C. Blum and A. Roli, "Metaheuristics in combinatorial optimization: Overview and conceptual comparison," *ACM Computing Surveys (CSUR),* vol. 35, no. 3, pp. 268-308, 2003.

[34] M.-W. Park and Y.-D. Kim, "A Systematic Procedre for Setting Parameters in Simulated Annealing Algorithms," *Computers Operations Research,* vol. 25, no. 3, pp. 207-217, 1998.

[35] P. van Laarhoven and E. Aarts, Simulated Annealing: Theory and Applications, Norwell: Kluwer Academic Publishers, 1987, pp. 7-15.

[36] M. Maghrebi and T. S. Waller, "Exploring Experts Decisions in Concrete Delivery Dispatching Systems Using Bayesian Network Learning techniques,"

in *7th European Modelling Symposium on Mathematical Modelling and Computer Simulation*, Manchester, 2013.

[37] R. Eglese, “Simulated Annealing : A Tool for Operational Research,” *European Journal of Operational Research,* vol. 46, pp. 271-281, 1990.

[38] J. McCall, “Genetic algorithms for modelling and optimisation,” *Journal of Computational and Applied Mathematics,* vol. 184, pp. 205-222, 2005.

[39] Q. Chen and C. Li, “An approach to bus-driver scheduling problem,” in *2010 Second WRI Global Congress on Intelligent Systems*, 2010.

[40] S. Sivanandam and S. Deepa, Introduction to Genetic Algorithms, Berlin: Springer, 2008.

[41] Y. Karuno and H. Nagamochi, “2-Approximation algorithms for the multi-vehicle scheduling problem on a path with release and handling times,” *Discrete Applied Mathematics,* vol. 129, pp. 433-447, 2003.

[42] “Google Distance Matrix API,” [Online]. Available: https://developers.google.com/maps/documentation/distance-matrix/.