

The Concept and Impact Analysis of a Flexible Mobility on Demand System

Bilge Atasoy ^{*} Takuro Ikeda [†] Moshe E. Ben-Akiva ^{*}

Abstract

This paper introduces an innovative transportation concept called Flexible Mobility on Demand (FMOD), which provides personalized services to passengers. FMOD is a demand responsive system in which a list of travel options is provided in real-time to each passenger request. The system provides passengers with flexibility to choose from a menu that is optimized in an assortment optimization framework. For operators, there is flexibility in terms of vehicle allocation to different service types: taxi, shared-taxi and mini-bus. The allocation of the available fleet to these three services is carried out dynamically and based on demand and supply so that vehicles can change roles during the day. The FMOD system is built based on a choice model and consumer surplus is taken into account in order to improve the passenger satisfaction. Furthermore, profits of the operators are expected to increase since the system adapts to changing demand patterns. In this paper, we introduce the concept of FMOD and present preliminary simulation results that quantify the added value of this system.

Keywords: demand responsive transit system, flexible transportation services, personalized transportation services, shared-taxi, shared-ride, dial-a-ride, mobility on demand

1 Introduction and motivation

Flexible forms of transportation are of interest to both passengers and transportation operators for many reasons. Many individuals, such as elderly and disabled passengers, have difficulty using conventional public transportation services. Additionally, conventional public services may be inconvenient since they are not personalized and may not meet the needs of the passengers intended trip. Public transportation services have fixed routes and schedules, and they may have low frequency during certain times of the day or in some parts of the transportation network. As a result, these services tend to have low ridership. Conversely, taxis tend to provide greater flexibility but they are unaffordable on a regular basis for many individuals.

^{*}Massachusetts Institute of Technology, {batasoy@mit.edu}, {mba@mit.edu}

[†]Fujitsu Laboratories, Ltd, {ikeda.takuro@jp.fujitsu.com}

From the perspective of transportation operators, in areas where the public transportation services have low ridership, profitability decreases due to under-utilization of the transportation capacity. Because of small profit margins, it is economically inefficient for transportation operators to provide higher frequency or better coverage. As a result they cannot offer a high quality of service. A vicious cycle is created where low ridership leads to lower profit for operators and decreased quality of service, which in turn results in lower passenger satisfaction and ridership.

In order to break this vicious cycle, new and innovative transportation alternatives must be developed. Several alternatives, such as dial-a-ride services, have been introduced around the world. These services are usually provided to elderly and handicapped people who cannot use regular public transportation alternatives. The need for such flexible services is expected to increase due to the aging population and the increased importance given to equity and quality of life issues. As flexible services are introduced, a critical issue will be how best to ensure the profitability of transport operators while providing a high quality of service. Due to the nature of the target population, the quality of service is highly important and increases operating costs.

In this paper, we introduce the Flexible Mobility on Demand (FMOD) system, which provides flexibility to both passengers and transportation operators. FMOD provides different levels of service to each passenger request; passengers are presented with a menu of choices from which they make a selection based on their preferences. The flexibility provided to the operators is based on the allocation of the vehicles to different services depending on the received request.

FMOD utilizes three services: taxi, shared-taxi and mini-bus. A list of travel options is designed in real-time and includes these three types of services. These services vary in regard to flexibility and cost. For instance, while taxi services are convenient for passengers because they provide door-to-door service at the passengers desired time, they are also expensive to operate and not an affordable mode of regular transportation for many passengers. Conversely, fixed-route mini-bus services can be cheaper to operate because passengers and origin-destination pairs can be pooled together. However, fixed routes and schedules may not be as convenient as taxi service, especially for passengers who are elderly, disabled, or traveling to outlying areas during off-peak times. The flexible nature of FMOD aims to improve the sustainability of transportation systems while simultaneously improving convenience for passengers and profitability for operators.

This paper introduces the concept of FMOD and provides preliminary analysis that

serves as a proof of concept. An optimization framework has been developed; when a request is received, the menu of choices offered is optimized by taking into account passenger preferences and operational constraints related to schedule and capacity. Preliminary results have been obtained for a network in Tokyo under different scenarios.

The rest of the paper is organized as follows: section 2 presents related literature and section 3 introduces the FMOD system. Section 4 presents the methodology for the operation of the FMOD system. Section 5 presents preliminary results and evaluation of the system based on simulation experiments. Finally, we conclude the paper in section 6.

2 Related literature

Demand Responsive Transit (DRT) has been increasingly studied and applied in the last decade. DRT is a user-oriented form of public transport with flexible routing and scheduling based on passenger needs. DRT is usually operated in a shared-ride mode between requested pick-up and drop-off locations. The most well known version of DRT is called *dial-a-ride* where door-to-door transportation services are provided using mini-buses and taxis. A fully automated version has been introduced by Dial (1995) and is called Autonomous Dial-A-Ride Transit (ADART).

The routing and scheduling of the vehicles in dial-a-ride systems is referred to as dial-a-ride problem (DARP), and it is addressed by various operations research techniques in the literature. We refer to the paper by Cordeau and Laporte (2007) for a comprehensive review of models and algorithms developed for DARP in the literature. In general, the optimization problem aims to find the set of minimum cost vehicle routes while accommodating as many requests as possible under a set of operational and quality of service constraints. We cite a few operations research approaches for the solution of DARP; Cordeau and Laporte (2003) address the problem with a tabu search heuristic, Coslovich et al. (2006) propose a two-phase insertion heuristic for the solution of the version with time windows, and Parragh et al. (2010) present a variable neighborhood search approach.

DARP is similar to a number of vehicle routing problems in the literature. We refer to Toth and Vigo (2001) for a review on the vehicle routing problems. The main distinction of DARP from vehicle routing problems is that incorporating the users perspective when providing a convenient service is an important objective along with minimizing the operating costs. The considered DRT systems are mostly studied or applied to a niche market for elderly or disabled people, especially in rural areas where the demand is typically low and spread over a large area.

More recently, the concept of DRT has been broadened to go beyond its niche market and is referred to as Flexible Transport Services (FTSs). Mulley and Nelson (2009) provide an overview of FTS stating that the scope is to improve the flexibility and convenience of public transport and to keep a comparable price to existing public transport services. Brake et al. (2007) analyze the recent experiences with FTS in the US and Europe where they come up with policy insights regarding the design, operation and the technology to be used in the context of FTS.

In this paper, we introduce a Flexible Mobility on Demand (FMOD) system that combines the above-mentioned concepts of flexible transport alternatives with an extended notion of flexibility. It is a reservation-based system which offers ride alternatives to passenger requests in real-time. The innovative aspect in the provided flexibility is that a list of options is offered to each request so that the passenger has multiple alternatives to choose from. This is considered an assortment optimization problem. Assortment optimization problems have been widely studied in operations research literature.

Talluri and van Ryzin (2004) present an airline revenue management model that optimizes the set of products to be offered at each point in time based on a general discrete choice model in an unconstrained setting. They formulate the model as a dynamic programming problem and characterize the optimal policy that maximizes the revenue. They show that the optimal assortment consists of a number of products with the largest revenues under a multinomial logit model (MNL).

Rusmevichientong et al. (2010) study the assortment optimization problem subject to a capacity constraint, which can be considered as the shelf space constraint observed by the retailers when they optimize the assortment of the products on the shelves. They represent the demand by MNL for homogenous customers. They present algorithms for both the static and dynamic versions of the problem where the former assumes that the parameters of the MNL are known in advance and the latter learns the parameters in an adaptive framework. Rusmevichientong et al. (2014) extend the framework with random parameters for MNL, namely mixture of MNL models, which represents the existence of several customer segments with different preferences in the population. They come up with tractable problems for a specific class of random choice parameters and provide approximations for the general case.

Gallego et al. (2011) present a linear programming formulation for the assortment problem under MNL where all possible combinations of the products are considered feasible. Davis et al. (2013a) study different assortment problems based on MNL and show

that these problems can all be solved as a linear program. The variants of the assortment optimization problems they study include the introduction of cardinality constraints, the concept of display location based attractiveness, the integration of the pricing decision from a list of possible price levels, the quality consistent pricing and the concept of product precedence constraints.

Davis et al. (2013b) study the assortment optimization problem, without any constraints on the products, under a nested logit model, where customers first choose a nest of products and then a product from the nest. They identify special cases where the problem is polynomially solvable and they develop tractable methods to obtain the assortments for the NP-hard cases of the problem. Gallego and Topaloglu (2013) extend the assortment problem under the nested logit model towards the introduction of cardinality and space constraints on the offered assortment. The constrained assortment optimization problem is also further extended in order to jointly decide on the offered assortment and the price of the products in the assortment.

The listed references for the assortment optimization rely on discrete choice models in order to represent the response of customers. In the context of mode choice models, there are few studies where flexible alternatives are considered in the choice set. Based on a survey conducted in the San Francisco Bay Area in 1990, several choice models are developed (Bhat, 2000; Koppelman and Bhat, 2006) including the alternative of shared-ride. More recently, Yang (2010) presents a master thesis with stated preferences (SP) survey methodologies and estimates mode choice models including the alternatives of regular taxi, shared-taxi, express mini-bus and others. The FMOD project in the long run includes data collection and demand model estimation for the proposed transportation alternatives. However, this is out of the scope of the current paper and the considered demand model in this paper is based on the literature.

As mentioned before, this paper introduces the concept of FMOD. In the future, we aim to conceptualize FMOD as a personalized mobility management system. The ideas behind the mobility management system are inspired by recently developed mobility applications and *recommender systems* that are used in several contexts. *Ridescout*¹ is a recently developed mobile application that provides real-time information on available modes together with the estimated cost and travel time for the requested origin-destination. Recommender systems, on the other hand, give recommendations to the users based on the ratings of a product. These ratings are estimated based on the historical ratings of the user on similar

¹<http://www.ridescoutapp.com>

products or ratings of users with similar tastes and preferences. For an overview of recommender systems we refer to Adomavicius and Tuzhilin (2005). The FMOD system will ultimately make use of similar methodologies. However FMODs main distinction is that the alternatives provided to each passenger request are optimized in real-time taking into account his/her preferences, rather than listing all potential alternatives. Furthermore, as done in recommender systems, FMOD will be designed such that as passengers use the services, the FMOD system will learn the behavior and the demand model will be calibrated. Therefore, the FMOD system will better serve passengers because of the ability to understand and incorporate their personal preferences.

3 Flexible Mobility on Demand (FMOD)

In this section we describe the FMOD system and the way the reservation system works.

3.1 FMOD Concept

FMOD integrates several methodologies including assortment optimization, vehicle routing, scheduling and pricing for the design of travel options in real-time. In this paper the main focus is the assortment optimization problem. As depicted in Figure 1, FMOD has flexibility in terms of the level of service provided by the vehicles in the fleet. The FMOD fleet is assumed to be a homogenous fleet consisting of vans that typically have a capacity of 8 seats. The FMOD system has different levels of services with the available fleet of vehicles depending on the evolving demand in the network. In other words, each vehicle changes its role dynamically among three types of services: taxi, shared-taxi, and mini-bus in the planning horizon.

Each level of service offered by FMOD is characterized as follows:

- **Taxi service (T)** serves a single passenger at a time and provides door-to-door service. Passengers can board and alight at arbitrary locations. Since this is a private door-to-door service it has the highest price.
- **Shared-taxi service (S)** serves multiple passengers in the same vehicle and provides door-to-door service. Though passengers can board and alight at arbitrary locations, travel time may increase due to the pick-up and drop-off of other passengers.
- **Mini-bus service (B)** serves similar to a regular bus service. It runs along fixed routes and passengers board / alight at pick-up / drop-off locations on the routes.

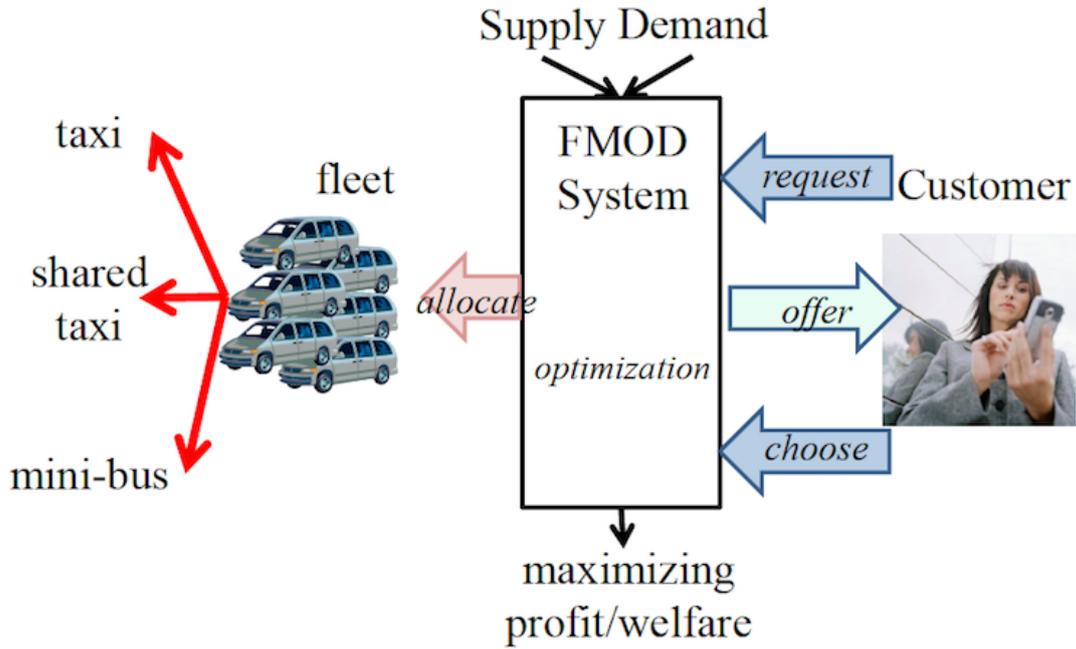


Figure 1: Flexible vehicle allocation based on passenger request

Pick-up / drop-off locations are predetermined bus stops or arbitrary points on the routes. The scheduling of the mini-bus service is not fixed but rather adapted to the preferred schedule of passengers similar to the shared-taxi service. Therefore, the demand responsiveness of the mini-bus is maintained through its flexible schedule. This implies that a vehicle starts serving as a mini-bus only when the first passenger is assigned to it as opposed to regular mini-bus services that run on a fixed schedule regardless of the demand.

Figure 2 shows the components of FMOD. A passenger requests a ride using a device such as smart phone, tablet or laptop. After the reservation is confirmed, one vehicle in the fleet will be notified of the schedule via on-board device or smart phone. Since each vehicle uploads GPS data continuously, FMOD can keep track of its location.

3.2 Reservation procedure

Figure 3 illustrates the reservation procedure of FMOD. In step 1, a passenger sends a ride request to the FMOD server. A ride request includes following information:

- Origin of the requested trip
- Destination of the requested trip
- Preferred departure time / Preferred arrival time

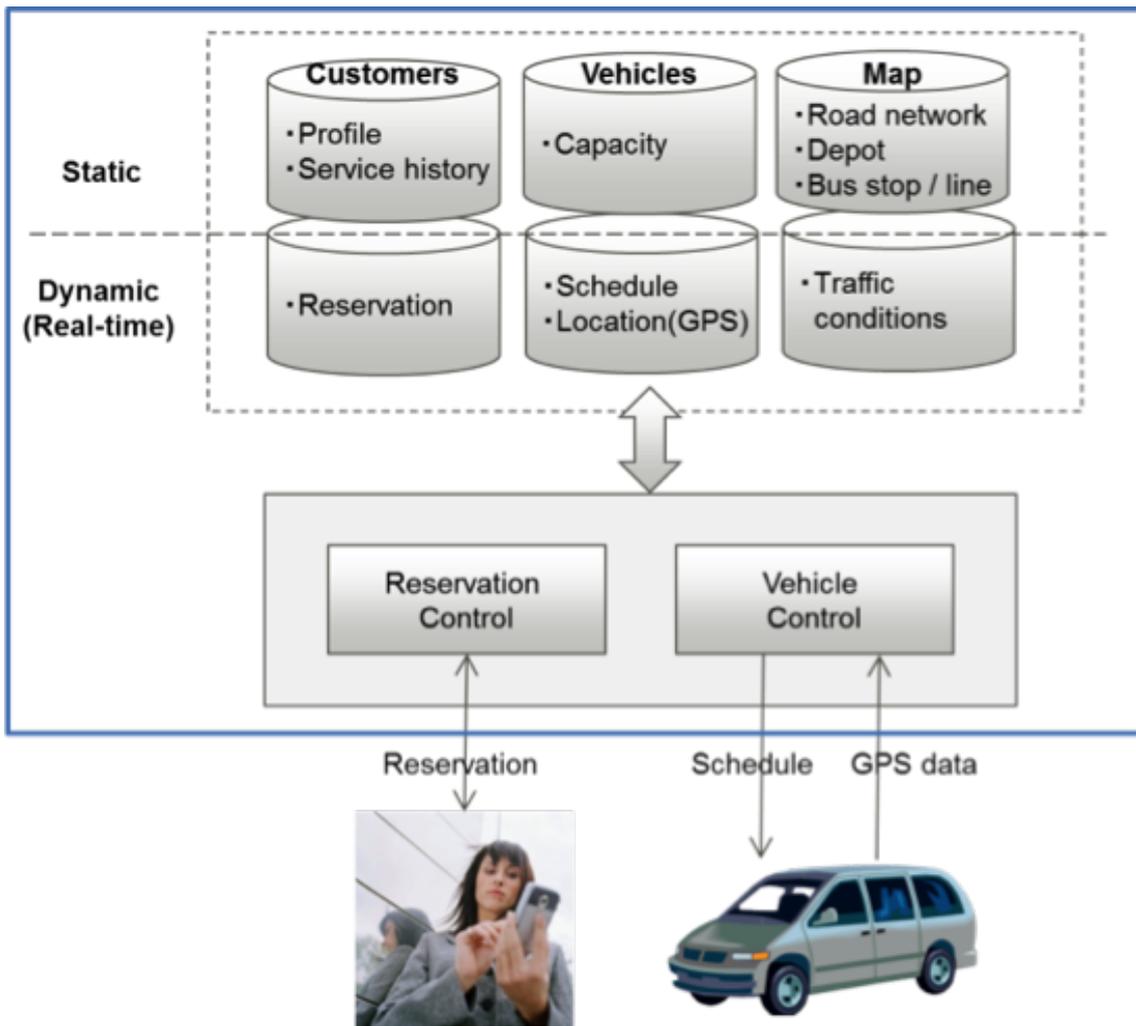


Figure 2: Components of the FMOD system

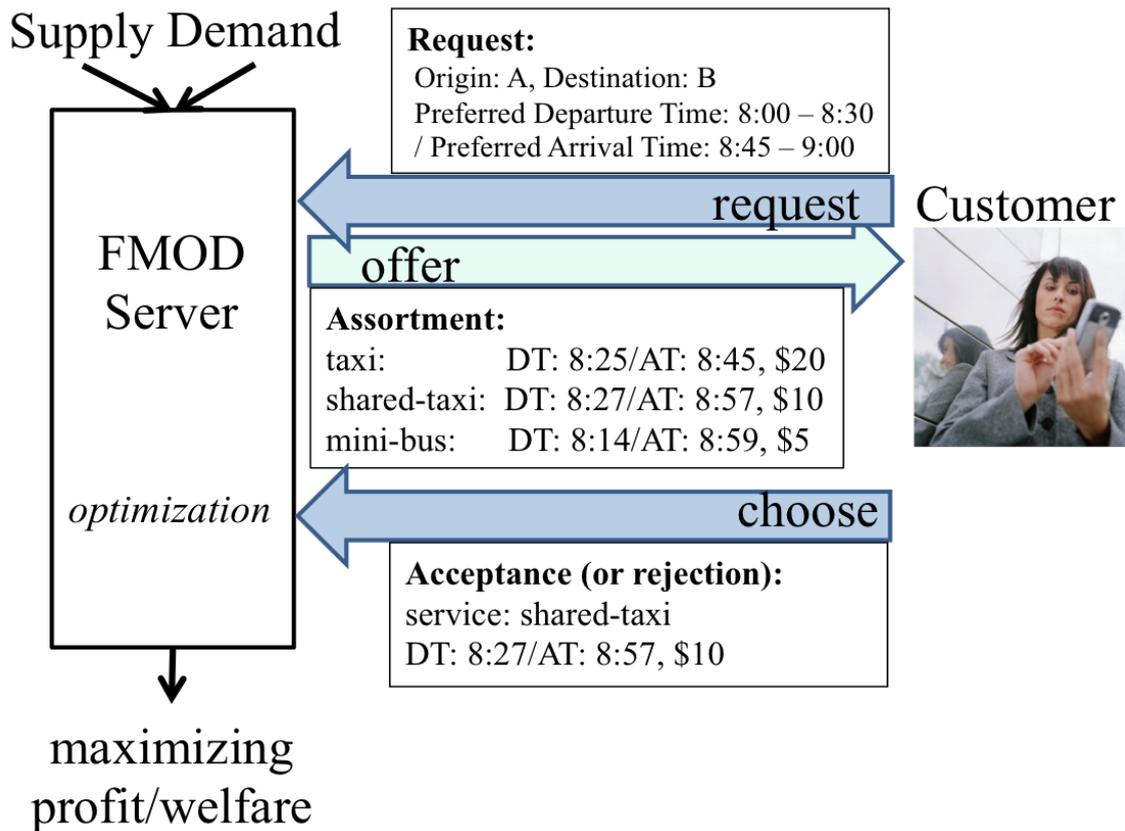


Figure 3: Reservation procedure

- Number of passengers

In terms of preferred departure / arrival time, the passenger can specify a time point or a time window when she or he wants to depart / arrive. The passenger can also request an immediate ride (c.f. "right now"). In step 2, the FMOD server generates a choice set which consist of ride alternatives for each service type, depending on the passenger request. The list of alternatives is presented to the passenger. Hereinafter the ride alternative is referred to as product and the choice set is referred to as assortment. Each product has following attributes:

- Service type (taxi, shared-taxi or mini-bus)
- Pick-up location
- Drop-off location
- Scheduled pick-up time
- Scheduled drop-off time

- Fare

For taxi and shared-taxi services, pick-up and drop-off locations are identical to the origin and destination of the requested trip respectively. For the mini-bus service, pick-up and drop-off locations are the nearest bus stops on bus routes.

In step 3, the passenger chooses one product out of the assortment and the server is notified. The passenger may reject all the products in the assortment meaning that he/she does not use the FMOD system. Finally in step 4, the FMOD server sends a confirmation to the passenger that the service is confirmed.

4 Modeling framework

As mentioned before, the novelty of the FMOD system is that the customer is presented a list of options for his/her requested trip in real-time. This assortment of products is selected based on an optimization procedure. The optimization procedure is handled in two phases:

Phase 1. Feasible product set generation

Phase 2. Choice-based assortment optimization

In phase 1, a set of feasible products is generated for each service type according to the customer request. A feasible product needs to satisfy the capacity and scheduling constraints. Therefore the FMOD system examines all the possible options in order to match the supply to the passenger request.

The set of feasible products generated in phase 1 is an input to phase 2, where the best assortment of products is formed out of the products in the feasible set. The assortment optimization model decides on the assortment that maximizes the profit based on a discrete choice model.

The optimization procedure is presented in more details in the remainder of this section.

4.1 Feasible product set generation

In this section we present the methodology for the feasible product set generation phase. Consider that N represents the fleet of vehicles and M represents the set of services. In the context of FMOD, we assume that the number of vehicles is exogenous, i.e. it is a constant input to the optimization procedure. As mentioned in section 3.1, we have three service types under set M : taxi, shared-taxi and mini-bus.

For each vehicle and service type there may be several potential products for different time periods. Consider that there is a maximum number of products for each vehicle-service pair denoted by L . Therefore each product can be represented by $p_{n,m,l}$ for vehicle $n \in N$, service $m \in M$ and for different schedule options $l \in L$. P represents the set of products, which is of size $N \times M \times L$. When a passenger request arrives, FMOD comes up with a feasible set of products F among the products in set P , as given below:

$$F \subseteq P = \{p_{n,m,l}\} \quad n \in N, m \in M, l \in L \quad (1)$$

The set of feasible products, F , should satisfy schedule and capacity related constraints, which will be defined and formulated in the following subsections.

4.1.1 Schedule definition

A product is associated with a route and schedule assigned to a vehicle. For the representation of the network, a convention similar to Fu (2002) is considered. Each vehicle has a sequence of schedule blocks (SB_1, SB_2, \dots, SB_K) each of which consist of a sequence of pick-up and drop-offs. K denotes the number of SB s assigned to the vehicle. Each SB is defined by a deadheading movement, i.e. the vehicle starts empty to pick-up the first passenger and ends empty after dropping off the last passenger.

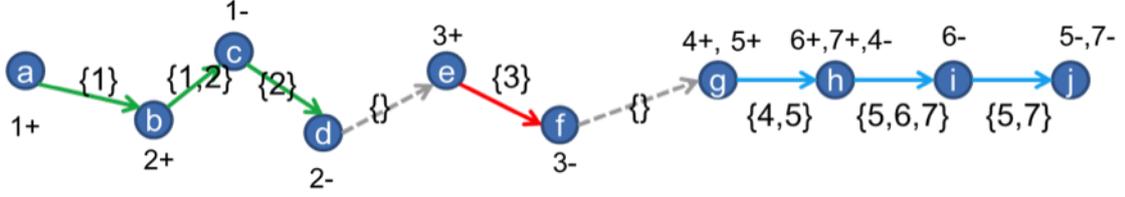
Each schedule block, SB_k , is associated with a sequence of stops that can be labeled as $(s_1^k, s_2^k, \dots, s_{S_k}^k)$, where S_k represents the number of stops for SB_k . Actually, each SB is associated with a vehicle and vehicles have different number of SB s in their sequence. However, for the ease of notation, we omit the index for the vehicles when referring to SB s and stops throughout the paper.

The SB s of each vehicle are associated with a service type meaning that the vehicle is assigned to a service or it is empty meaning that it is not serving any passenger. A SB is created when the first passenger is assigned to the vehicle and it is accomplished when all the passengers alight. When a SB is finalized, the vehicle moves from the final destination to the origin of the next SB under an empty schedule. A stop schedule includes the following information:

- Stop location
- Scheduled arrival time
- Scheduled departure time

[Schedule Block]	SB_1				SB_2		SB_3		SB_4		SB_5			
Service type	Shared				empty		taxi		empty		Mini-bus			
[Stops]	s_1^1	s_2^1	s_3^1	s_4^1	s_1^2	s_2^2	s_1^3	s_2^3	s_1^4	s_2^4	s_1^5	s_2^5	s_3^5	s_4^5
Location	a	b	c	d	d	e	e	f	f	g	g	h	i	j
Arrival Time	-	9:15	9:30	9:50	-	13:00	-	13:30	-	14:40	-	15:00	15:10	15:20
Departure time	9:00	9:16	9:31	-	12:30	-	13:05	-	14:30	-	14:50	15:01	15:11	-
Boarding passengers	1	2	-	-	-	-	3	-	-	-	4,5	6,7	-	-
Alighting passengers	-	-	1	2	-	-	-	3	-	-	-	4	6	5,7

(a) Schedule blocks



(b) Node representation

Figure 4: Demonstration of the schedule

- Boarding passengers
- Alighting passengers

Figure 4(a) shows an example of a sequence of SB s assigned to a vehicle. The vehicle first serves as a shared-taxi, then is empty until the taxi service and finally serves as a mini-bus after a relatively shorter empty time. Figure 4(b) demonstrates the movement of the vehicle according to the sequence. Nodes denote stop locations and the numbers associated with the nodes represent the boarding and alighting passengers. The numbers index the passengers, a + sign means that the passenger is boarding and a - sign indicates that the passenger is alighting. Furthermore, the lists associated with the links give the list of passengers that are on board between two stops.

The sequence of SB s is updated upon a new passenger request in two ways:

Case 1: A new schedule block SB^* is created and inserted into the sequence of existing SB s. The existing sequence is updated as follows:

$$\begin{aligned} \exists k(SB_1, \dots, SB_k, SB_{k+1}, \dots, SB_K) &\Rightarrow (SB_1, \dots, SB_k, SB^*, SB_{k+1}, \dots, SB_K) \\ &\Rightarrow (SB_1, \dots, SB_{K+1}) \end{aligned} \quad (2)$$

This happens when a new role is assigned to the vehicle, which means the new request cannot simply be served by its current role. For the inclusion of the new SB , shortest path algorithm is exploited for taxi and shared-taxi since they do not have pre-determined routes. However, for the mini-bus service we select the shortest route among the pre-determined routes taking into account the accessibility of bus stops.

Case 2: A pair of stops (pick-up and drop-off) is inserted into one of the existing schedule blocks, SB_k . Note that the new pick-up and drop-off stops might already be included in the network due to previous requests. Updated SB is referred as SB_k^* and the new stops are denoted by s'^k and s''^k . The modification in the sequence of SB s is given in (3) and the update of the sequence of stops is given in (4).

$$\exists k(SB_1, \dots, SB_k, \dots, SB_K) \Rightarrow (SB_1, \dots, SB_k^*, \dots, SB_K) \Rightarrow (SB_1, \dots, SB_K) \quad (3)$$

$$\begin{aligned} & \exists i, j(s_1^k, \dots, s_i^k, s_{i+1}^k, \dots, s_j^k, s_{j+1}^k, \dots, s_{S_k}^k) \\ & \Rightarrow (s_1^k, \dots, s_i^k, s'^k, s_{i+1}^k, \dots, s_j^k, s''^k, s_{j+1}^k, \dots, s_{S_k}^k) \Rightarrow (s_1^k, \dots, s_{S_k+2}^k) \end{aligned} \quad (4)$$

This scenario happens when the new request can be served with one of the already assigned services of the vehicle so that a new stop is added to its existing route. For the insertion of new stops on the route, we apply the insertion algorithm that is commonly used for DARP in literature (e.g. Coslovich et al., 2006) due to its computational efficiency. For shared-taxi the insertion algorithm can be applied. However, for the mini-bus service, such an insertion is carried out if it is feasible to serve the new passenger with the already selected route based on the accessibility of bus stops. Since the taxi is a private mode, no insertion can be applied to the existing taxi service.

Figure 5 illustrates the schedule update upon a new ride request from a passenger. Case 1 is demonstrated by Figure 5(a) where a new schedule block, SB_3 , is generated for the vehicle as a taxi service. Case 2 is illustrated in Figure 5(b), where new stops at locations e and f are inserted into the existing schedule block, SB_1 , under the shared-taxi service.

4.1.2 Scheduling and capacity constraints

In this section we introduce the constraints that need to be satisfied for each feasible product. These constraints are related to the seat capacity of the vehicles and the requested departure/arrival time of passengers.

Seat capacity constraints

Consider that NP denotes the number of passengers on board of a vehicle between two consecutive stops. The number of passengers on board a vehicle cannot be more than the

[Schedule Block]	SB₁				SB₂		SB₃					
Service type	Shared				empty		Mini-bus					
[Stops]	s_1^1	s_2^1	s_3^1	s_4^1	s_1^2		s_2^2		s_1^3	s_2^3	s_3^3	s_4^3
Location	a	b	c	d	d		g		g	h	i	j
Arrival Time	-	9:15	9:30	9:50	-		14:40		-	15:00	15:10	15:20
Departure time	9:00	9:16	9:31	-	12:30		-		14:50	15:01	15:11	-
Boarding passengers	1	2	-	-	-		-		4,5	6,7	-	-
Alighting passengers	-	-	1	2	-		-		-	4	6	5,7



[Schedule Block]	SB₁				SB₂		SB₃		SB₄		SB₅			
Service type	Shared				empty		taxi		empty		Mini-bus			
[Stops]	s_1^1	s_2^1	s_3^1	s_4^1	s_1^2	s_2^2	s_1^3	s_2^3	s_1^4	s_2^4	s_1^5	s_2^5	s_3^5	s_4^5
Location	a	b	c	d	d	e	e	f	f	g	g	h	i	j
Arrival Time	-	9:15	9:30	9:50	-	13:00	-	13:30	-	14:40	-	15:00	15:10	15:20
Departure time	9:00	9:16	9:31	-	12:30	-	13:05	-	14:30	-	14:50	15:01	15:11	-
Boarding passengers	1	2	-	-	-	-	3	-	-	-	4,5	6,7	-	-
Alighting passengers	-	-	1	2	-	-	-	3	-	-	-	4	6	5,7

(a) Case 1

[Schedule Block]	SB₁				SB₂		SB₃					
Service type	Shared				empty		Mini-bus					
[Stops]	s_1^1	s_2^1	s_3^1	s_4^1	s_1^2		s_2^2		s_1^3	s_2^3	s_3^3	s_4^3
Location	a	b	c	d	d		g		g	h	i	j
Arrival Time	-	9:15	9:30	9:50	-		14:40		-	15:00	15:10	15:20
Departure time	9:00	9:16	9:31	-	12:30		-		14:50	15:01	15:11	-
Boarding passengers	1	2	-	-	-		-		4,5	6,7	-	-
Alighting passengers	-	-	1	2	-		-		-	4	6	5,7



[Schedule Block]	SB₁						SB₂		SB₃					
Service type	Shared						empty		Mini-bus					
[Stops]	s_1^1	s_2^1	s_3^1	s_4^1	s_5^1	s_6^1	s_1^2		s_2^2		s_1^3	s_2^3	s_3^3	s_4^3
Location	a	b	c	e	d	f	d		g		g	h	i	j
Arrival Time	-	9:15	9:30	9:40	9:50	10:20	-		14:40		-	15:00	15:10	15:20
Departure time	9:00	9:16	9:31	9:41	-	-	12:30		-		14:50	15:01	15:11	-
Boarding passengers	1	2	-	3	-	-	-		-		4,5	6,7	-	-
Alighting passengers	-	-	1	-	2	3	-		-		-	4	6	5,7

(b) Case 2

Figure 5: Demonstration for the update of the schedule blocks and stops

seat capacity of the service assigned to the vehicle. This constraint can be written for any consecutive stops s_i^k and s_{i+1}^k of SB_k as follows:

$$NP_{s_i^k, s_{i+1}^k} \leq \text{Seats}_m \quad \forall i \in (1, 2, \dots, S_k - 1), \quad (5)$$

where Seats_m is defined as the total seat capacity of the vehicle for service m . Recall that the fleet consists of homogenous type of vehicles but the capacity may change for different services.

Consecutive schedule block constraints

Let $B_{s_i^k}$ denote the location of each stop s_i^k in SB_k . Furthermore, $AT_{s_i^k}$ and $DT_{s_i^k}$ represent the arrival time to stop s_i^k and departure time from stop s_i^k , respectively. In order to have a feasible product, the sequence of SB s assigned to a vehicle must satisfy the following constraints:

$$B_{s_{S_k}^k} = B_{s_1^{k+1}} \quad \forall k \in (1, 2, \dots, K - 1) \quad (6)$$

$$AT_{s_{S_k}^k} \leq DT_{s_1^{k+1}} \quad \forall k \in (1, 2, \dots, K - 1) \quad (7)$$

Constraints (6) guarantee that there is no conflict in terms of the location between consecutive schedule blocks, SB_k and SB_{k+1} . The location of the last stop in SB_k is the same location as the first stop in SB_{k+1} . Note that the last stop in SB_k is represented by $s_{S_k}^k$ and the first stop in SB_{k+1} is given by s_1^{k+1} . If constraint (6) is not satisfied for a given vehicle, it means that the vehicle will be idle and an empty schedule block should be inserted between these two SB s. With this insertion all the schedule blocks will satisfy the constraints. Similarly, constraints (7) maintain the feasibility with respect to time; namely the arrival time at the last stop of SB_k should be earlier than the departure time from the first stop of SB_{k+1} .

Committed departure/arrival time constraints

When a passenger chooses one of the products in the assortment and once this reservation is confirmed by the system, scheduled departure and arrival times of the product will be referenced as *committed departure time* and *committed arrival time*, respectively. The schedule at the time of booking can be modified later based on the requests of the other passengers in the case of a shared-ride. Therefore, scheduled departure and arrival times can deviate from the committed departure and arrival times, respectively. However, for the sake of convenience to the passengers, we need to limit this change on the schedule.

Let CDT_p , SDT_p , CAT_p and SAT_p represent the committed departure time, scheduled departure time, committed arrival time and scheduled arrival time of a product that will serve a passenger. As mentioned before, the ride request can either be based on a preferred departure time or a preferred arrival time. For a passenger who specifies a preferred departure time, constraint (8) should be satisfied in order to have a feasible product. Similarly, for a passenger who specifies preferred arrival time, we must satisfy constraint (9). These constraints assume that there is a maximum allowance for deviation denoted by T^{max} . The SB that includes the feasible product p should follow this maximum allowance so that the schedule is built accordingly.

$$|CDT_p - SDT_p| \leq T^{max} \quad \forall p \in F \quad (8)$$

$$|CAT_p - SAT_p| \leq T^{max} \quad \forall p \in F \quad (9)$$

In a shared-taxi service, the departure and arrival time of the passenger changes, and the travel time may also increase due to the pick-up and drop-off of other passengers. For the sake of passenger convenience, this detouring should be controlled. $IVTT_p$ is defined as the in-vehicle travel time for each product p . Constraint (10) defines an upper limit, $IVTT_p^{max}$, on the in-vehicle travel time for each product. This upper limit can be decided based on the shortest travel time for the product in the absence of other passengers sharing the ride. This shortest travel time can be increased by a predetermined percentage to come up with $IVTT_p^{max}$.

$$IVTT_p \leq IVTT_p^{max} \quad \forall p \in F \quad (10)$$

4.1.3 Tight/Loose feasible products

A product is feasible if the schedule satisfies the constraints given in (5)-(10) for a given passenger request. Otherwise, the product is infeasible and cannot be included in the assortment. Furthermore we consider *tight-feasible* and *loose-feasible* products based on the preferred time windows of the passengers. We denote the set of tight-feasible products by F^{tight} and the set of loose-feasible products by F^{loose} where $F^{\text{tight}} \cup F^{\text{loose}} = F$.

Recall that SDT_p and SAT_p define the scheduled departure time and arrival time of a product. As mentioned in section 3.2, the ride request is either initiated with a departure time window or an arrival time window. We denote the departure and arrival time windows by (PDT^-, PDT^+) and (PAT^-, PAT^+) , respectively. If the ride request is initiated based on the departure time, a tight-feasible product must satisfy constraint (11). Similarly, for a passenger who specifies a preferred arrival time window, a tight-feasible

product must satisfy constraint (12).

$$PDT^- \leq SDT_p \leq PDT^+ \quad \forall p \in F^{\text{tight}} \quad (11)$$

$$PAT^- \leq SAT_p \leq PAT^+ \quad \forall p \in F^{\text{tight}} \quad (12)$$

For each request, the system generates a single tight-feasible product for each vehicle-service pair in the feasible set. However, there may be multiple loose-feasible products in the set according to the availability of the fleet. The deviation of the loose-feasible products from the preferred time window of passenger, namely the schedule delay, is defined by SD^{max} . In other words, the loose-feasible products can be outside the preferred time windows by at most SD^{max} time units. Therefore, the departure/arrival time for the feasible products including the loose-feasible ones are as follows:

$$PDT^- - SD^{\text{max}} \leq SDT_p \leq PDT^+ + SD^{\text{max}} \quad \forall p \in F \quad (13)$$

$$PAT^- - SD^{\text{max}} \leq SAT_p \leq PAT^+ + SD^{\text{max}} \quad \forall p \in F \quad (14)$$

Given these extended time windows for departure/arrival time, the set of loose-feasible products are generated as a result of two mechanisms. The first mechanism divides the time horizon for loose-feasible products given by (13) or (14) into a number of time intervals. Let us denote the length of these intervals by T^{interval} . For each time interval, the feasibility of offering the product at the beginning of the interval is evaluated. The maximum number of loose-feasible products that are either earlier or later than the preferred time window is given by $2(SD^{\text{max}}/T^{\text{interval}})$.

The second mechanism considers the vehicles that are already under service with some available capacity on board. For the time period defined by (13) or (14), the algorithm checks the feasibility of these vehicles and includes them in the set of loose-feasible products if feasible. Note that vehicles that are assigned to a taxi service will never be available since this service is a private mode. The estimation of the maximum number of feasible products in this second mechanism is not straightforward. We assume a maximum number of such products for the ease of implementation.

The defined mechanisms generate loose-feasible products in addition to the tight-feasible products as candidates to be offered to the passengers. The set of all feasible products is an input to the assortment optimization problem which is explained in the next section. During off-peak hours when demand is relatively low, the FMOD system will mostly select tight-feasible products to offer to the passengers since the choice probability is higher for those products. However, during peak hours, loose-feasible products might be offered more frequently due to limited capacity.

4.2 Choice-based assortment optimization

The methodology to generate a set of feasible products, F , is described in section 4.1. In this section we present the assortment optimization problem where the best assortment of products is selected among the set of feasible products. The objective of the assortment optimization problem is considered as maximizing the profit. As an initial step to show the proof of concept we focus on a myopic approach where we optimize the assortment based on the current ride request. Therefore, when a request is received the best assortment is generated among the feasible set of products regardless of the future possible requests.

We define a binary decision variable, $x_{n,m,l}$ for each feasible product $p_{n,m,l} \in F$. If the feasible product is included in the assortment $x_{n,m,l} = 1$, and it is 0 otherwise. The vector of x variables defines the choice set of the passenger. Therefore, the assortment optimization problem optimizes the choice set which should take into account the preferences of the passengers. A discrete choice model is integrated into the optimization framework in order to incorporate the passengers viewpoints.

In the remainder of this section, we introduce the discrete choice model and present the assortment optimization model.

4.2.1 Choice model

We assume that passengers make their choices among the set of products in the assortment based on a multinomial logit model (MNL). The choice set is defined by the decision variables $x_{n,m,l}$, so that the logit model is an endogenous component of the optimization model. Therefore the choice probabilities for the provided services are variables of the model rather than fixed inputs. As mentioned in section 3.2, the passengers have an additional alternative to reject all the offered services.

The deterministic part of the utility function for each product $p_{n,m,l} \in F$ is denoted by $V_{n,m,l}$. It is defined for the taxi, shared-taxi, mini-bus services and the reject option as

given in the following equations.

$$\begin{aligned}
V_{n,\text{taxi},l} = & \text{ASC}_{\text{taxi}} - \text{price}_{n,\text{taxi},l} - \text{VOT}^{\text{IVTT}} \cdot \text{IVTT}_{n,\text{taxi},l} \\
& - \text{VOT}^{\text{SDE}} \cdot \text{SDE}_{n,\text{taxi},l} - \text{VOT}^{\text{SDL}} \cdot \text{SDL}_{n,\text{taxi},l}
\end{aligned} \tag{15}$$

$$\begin{aligned}
V_{n,\text{shared},l} = & \text{ASC}_{\text{shared}} - \text{price}_{n,\text{shared},l} - \text{VOT}^{\text{IVTT}} \cdot (\text{IVTT}_{n,\text{shared},l} + \Delta\text{IVTT}_{n,\text{shared},l}) \\
& - \text{VOT}^{\text{SDE}} \cdot \text{SDE}_{n,\text{shared},l} - \text{VOT}^{\text{SDL}} \cdot \text{SDL}_{n,\text{shared},l}
\end{aligned} \tag{16}$$

$$\begin{aligned}
V_{n,\text{bus},l} = & \text{ASC}_{\text{bus}} - \text{price}_{n,\text{bus},l} - \text{VOT}^{\text{IVTT}} \cdot \text{IVTT}_{n,\text{bus},l} \\
& - \text{VOT}^{\text{SDE}} \cdot \text{SDE}_{n,\text{bus},l} - \text{VOT}^{\text{SDL}} \cdot \text{SDL}_{n,\text{bus},l} \\
& - \text{VOT}^{\text{OVT}} \cdot \text{OVT}_{n,\text{bus},l}
\end{aligned} \tag{17}$$

$$V_{\text{reject}} = \beta_{\text{dist}} \cdot \text{STD} \tag{18}$$

- ASC_{taxi} , $\text{ASC}_{\text{shared}}$ and ASC_{bus} are the alternative specific constants for the services of taxi, shared-taxi and mini-bus offered by the FMOD system, respectively. The reject option is considered as the reference and its constant is fixed to zero.
- The utility function normalized such that it is in monetary units (\$). The normalization can be considered as dividing the utility function by the price parameter ($\beta_{\text{price}} < 0$). Therefore, the parameters for the travel time variables represent the value of time (VOT). Travel time variables include the in-vehicle travel time (IVTT), early schedule delay (SDE) and late schedule delay (SDL). For the shared-taxi service, there may be additional in-vehicle travel time due to other passengers sharing the ride. Therefore, ΔIVTT is defined for this additional travel time. For the mini-bus, the out-vehicle time (OVT) is the sum of access and egress time for passengers.
- The utility for the reject option should represent the available alternatives other than FMOD. As the travel distance increases, the decrease in the utility should be reflected. Therefore, the shortest travel distance (STD) for the requested trip is included as an explanatory variable. The associated parameter is denoted by β_{dist} . Since the utilities are in monetary units, it represents the monetary value of travel distance.

For the disutility associated with schedule delay, we follow the idea proposed by De Palma et al. (1983) and Ben-Akiva et al. (1986). As defined in section 4.1.3, a tight-feasible product is always in the preferred departure or arrival time windows which are given by (PDT^-, PDT^+) and (PAT^-, PAT^+) , respectively. Therefore, SDE and SDL

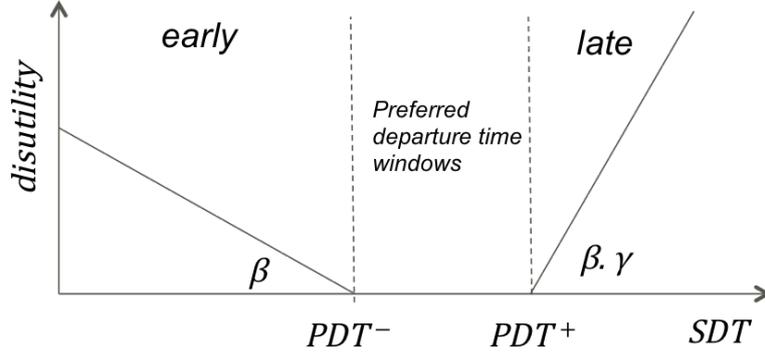


Figure 6: Disutility of schedule delay

will be zero for a tight-feasible product. On the other hand, a loose-feasible product departs/arrives either earlier or later than the preferred time window. It is assumed that the disutility linearly increases with respect to the delay. However, the impact of late schedule delay is higher than the impact of early schedule delay. The SDE and SDL are defined for the requests specified with a departure time by (19). Similarly, for the requests initiated with an arrival time window, they are given by (20). Figure 6 illustrates the applied idea with the deviation from the preferred time window. With the assumption of $\beta > 0$ and $\gamma > 1$, it is ensured that being late is penalized more than being early with respect to the preferred time window. Therefore, for our case it is assumed that $\text{VOT}^{\text{SDL}} = \gamma \text{VOT}^{\text{SDE}}$.

$$\begin{aligned} \text{SDE}_p &= \max(0, (\text{PDT}^- - \text{SDT}_p)) \\ \text{SDL}_p &= \max(0, (\text{SDT}_p - \text{PDT}^+)) \quad \forall p \in F \end{aligned} \quad (19)$$

$$\begin{aligned} \text{SDE}_p &= \max(0, (\text{PAT}^- - \text{SAT}_p)) \\ \text{SDL}_p &= \max(0, (\text{SAT}_p - \text{PAT}^+)) \quad \forall p \in F \end{aligned} \quad (20)$$

Given the utility functions, the choice probability for a feasible product $p_{n,m,l} \in F$ is given by:

$$\text{Prob}_{n,m,l}(x) = \frac{x_{n,m,l} \exp(\mu V_{n,m,l})}{\exp(\mu V_{\text{reject}}) + \sum_{n' \in N} \sum_{m' \in M} \sum_{l' \in L} x_{n',m',l'} \exp(\mu V_{n',m',l'})}, \quad (21)$$

where μ is the scale parameter that needs to be adjusted to $-\beta_{\text{price}}$ since we work with normalized utilities. Since the choice set will be determined by the optimization model through the decision variables x , the choice probability is a function of x variables.

4.2.2 Assortment optimization model

In this section we present the assortment optimization model that is integrated with the choice model presented in section 4.2.1. The objective function is specified in order to maximize the profit obtained by the operator.

The profit associated with a product $p_{n,m,l}$ is denoted by $r_{n,m,l}$, which is obtained as the difference between the price of the service and the operating costs. We present the optimization problem as follows:

$$\max \sum_{n \in N} \sum_{m \in M} \sum_{l \in L} r_{n,m,l} \text{Prob}_{n,m,l}(x) \quad (22)$$

$$\text{s.t.} \quad \sum_{n \in N} \sum_{l \in L} x_{n,m,l} \leq 1 \quad \forall m \in M \quad (23)$$

$$x_{n,m,l} \leq 0 \quad \forall p_{n,m,l} \notin F, n \in N, m \in M, l \in L \quad (24)$$

$$x_{n,m,l} \in \{0, 1\} \quad \forall n \in N, m \in M, l \in L \quad (25)$$

The objective function is the expected profit obtained from the list of offered alternatives as given by (22). Constraints (23) maintain that there will be at most one product for each service in the assortment. In the feasible set, there may be several alternatives for each service with different vehicles and time slots. However, as a result of the optimization there needs to be at most one option for each service. Constraints (24) ensure that x variables are 0 for products that are not in the feasible set F . Finally, we have the definition of binary x variables in (25).

The presented model in (22)-(25) is a mixed integer nonlinear problem. The nonlinearity is due to the probability term which is a function of the x variables given by (21). As Davis et al. (2013a) presented, this problem can be represented by a linear programming problem with a simple transformation. A new set of decision variables, w , is introduced that represents the choice probability. $\omega_{n,m,l}$ will be zero when the product $p_{n,m,l}$ is not offered and it will be positive and ≤ 1 for the offered products. ω_{reject} is similarly the probability of rejecting all the products in the assortment. The linear programming formulation is given below:

$$\max \sum_{n \in N} \sum_{m \in M} \sum_{l \in L} r_{n,m,l} \omega_{n,m,l} \quad (26)$$

$$\text{s.t.} \quad \sum_{n \in N} \sum_{m \in M} \sum_{l \in L} \omega_{n,m,l} = 1 - \omega_{reject} \quad (27)$$

$$\sum_{n \in N} \sum_{l \in L} \frac{\omega_{n,m,l}}{\exp(\mu V_{n,m,l})} \leq \frac{\omega_{reject}}{\exp(\mu V_{reject})} \quad \forall m \in M \quad (28)$$

$$\omega_{n,m,l} \leq 0 \quad \forall p_{n,m,l} \notin F, n \in N, m \in M, l \in L \quad (29)$$

$$0 \leq \frac{\omega_{n,m,l}}{\exp(\mu V_{n,m,l})} \leq \frac{\omega_{reject}}{\exp(\mu V_{reject})} \quad \forall n \in N, m \in M, l \in L \quad (30)$$

Constraint (27) means that the probabilities sum up to 1 including the reject option. Constraints (28) maintain that only one product is offered for each service analogous to constraints (23) in the original formulation. Constraints (29) ensure that infeasible products are not included in the assortment. The transformed model considers the relative attractiveness of the products such that the resulting choice probability should be proportional to the attractiveness of the product. In other words, ratio of $\frac{\omega_{n,m,l}}{\exp(\mu V_{n,m,l})}$ is controlled with respect to the ratio for the reject option $\frac{\omega_{reject}}{\exp(\mu V_{reject})}$. This relation is given by constraints (30).

5 Simulation experiments

In order to quantify the added value of FMOD, several experiments are conducted using a simulation framework where the models and methodologies described in section 4 are brought together. The whole framework is implemented in C++ and the assortment optimization problem (26)-(30) is solved by the LP solver provided in R. The time horizon is considered as 24 hours for the simulation. The simulation ignores the actual traffic conditions on the network.

As a case study we consider the Hino city in Tokyo which has a land area of approximately 9km \times 8km. The map of the city captured by Open Street Map data is given in Figure 7. The network consists of 31,287 nodes and 63,463 links.

In the remainder of this section we first provide the assumptions on the supply and demand model parameters and then we present experimental results.

5.1 Supply model parameters

The supply model parameters are related to the capacity and scheduling of the vehicles and the revenue and operating cost for the transport operator. As mentioned before, we



Figure 7: Map of Hino city

base our analysis on the three services provided by FMOD: taxi, shared-taxi and mini-bus. For this analysis we assume that there are 60 vehicles in the fleet which dynamically change their service role during the simulation. Each vehicle is considered to have 8 seats in total. However the capacity for the taxi service is 1 passenger only. Therefore, the capacity of each vehicle related to the set of constraints (5) is defined as follows:

$$\begin{aligned} \text{Seats}_{\text{taxi}} &= 1 \\ \text{Seats}_{\text{shared-taxi}} &= \text{Seats}_{\text{mini-bus}} = 8 \end{aligned}$$

The scheduling and assignment of drivers are not considered. It is assumed that there is an available driver for each allocated vehicle. Since we ignore the traffic conditions on the network, the assigned vehicles are assumed to follow their schedules without any delay due to traffic. The maximum allowance for the deviation from the committed schedule, T^{\max} , that is given in constraints (8) and (9) is assumed to be 10 minutes. This means that the committed departure/arrival times can be changed by at most 10 minutes due to new arriving passengers who will share the ride. This change in the committed schedule is also limited by $IVTT_p^{\max}$ for the in-vehicle time as given in constraints (10). $IVTT_p^{\max}$ is assumed to be twice the time of the taxi service for that particular request.

As mentioned in section 3.1, the mini-bus service operates between predefined bus stops rather than serving between the preferred origin and destination locations. The access and egress time for mini-bus is calculated based on the shortest distance between the bus stop and the origin/destination of the trip. A walking speed of 80 m/min is considered as it is well accepted as a preferred walking speed. If the origin/destination is more than 2 km away from the nearest bus stop, FMOD does not offer such a mini-bus service since it is considered to be inconvenient for passengers.

As given in section 4.1.3, the loose-feasible products are defined by the maximum schedule delay, SD^{\max} , with respect to the preferred departure/arrival time as given in constraints (13) and (14). SD^{\max} is assumed to be 90 minutes and the time interval for the generation of loose-feasible products, T^{interval} , is assumed to be 15 minutes. Therefore, with the first mechanism presented in section 4.1.3 there can be maximum 12 loose-feasible products generated. In addition, we include all the feasible products generated by the second mechanism in the feasible set. In other words, the already running vehicles with available capacity on board are included in the feasible set if their schedules satisfy the constraints.

The price for using the services are given in Table 1. It is assumed that the taxi service has a base price which is the same for any request and there is a variable price depending on the distance of the trip. The price for shared-taxi is considered to be half of taxi price. The mini-bus service has a flat rate of \$ 3 for each trip. Finally, the operating cost of the system is divided into fixed cost and variable cost. Fixed cost is assumed to be \$ 200 per day per vehicle and variable cost is assumed to be \$ 0.2 per km. Therefore, the profit $r_{n,m,l}$ for each product $p_{n,m,l}$ is given by the price of the product minus the variable operating cost.

Table 1: Price for FMOD services

service	price
Taxi	\$ 5 (base) + \$ 0.5 (per 320 m)
Shared-taxi	\$ 2.5 (base) + \$ 0.25 (per 320 m)
Mini-bus	\$ 3

5.2 Demand model parameters

The demand related parameters are of two types: the daily demand for FMOD and the logit model parameters which define the preferences of passengers towards the attributes of the alternatives.

It is assumed that there are 5000 ride requests in a day and the time-of-day distribution of the daily demand is given in Figure 8. The assumed demand pattern reflects the fluctuations during the day due to peak and off-peak hours. This demand variation is considered when generating the preferred departure time window such that we have a higher probability of receiving requests with the preferred departure time during peak hours. We note that the system is flexible to handle both preferred departure and arrival time windows, but in the simulation we assume that all requests are initiated with a preferred departure time window of 30 minutes.

The origin and destination of the requested trip is arbitrarily assigned in the area based on the population density meaning that more requests are generated from the areas with higher population density. Origins and destinations that are less than 500 meters apart are not included as a trip in the simulation since it is assumed to be a too short distance in order to request a service. The origin and destination locations can be any point of interest such as a city hall, hospital, train and bus stations.

Since FMOD is a reservation-based system, it is important to consider how long in advance passengers initiate their requests. In our simulation framework, we assume that

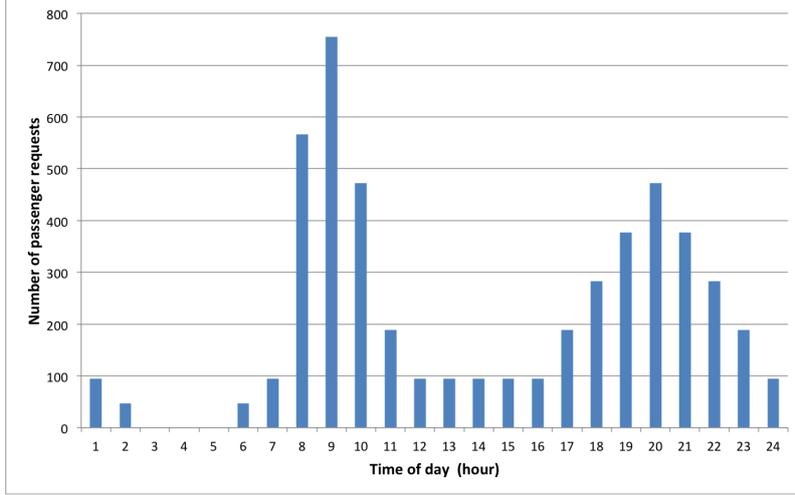


Figure 8: Time of day distribution of daily demand

the time between the request and the center of the preferred time window is normally distributed with mean and standard deviation of one hour. The normal distribution assumption here may generate preferred departure time windows earlier than the request time, which is not meaningful. Therefore, the normal distribution is truncated so that such cases are discarded and not included in the simulation.

The logit model introduced in section 4.2.1 represents the preferences of passengers towards the list of travel options presented by the FMOD system. The coefficients of the logit model are assigned based on the literature. First, the scale of the model is given by the $\mu = -\beta_{\text{price}}$ and it is assigned to be 0.5 based on the price parameters estimated in the literature (e.g. a study in San Francisco Bay Area by Koppelman and Bhat, 2006).

The alternative specific constants for different services are assigned based on our experiments and intuition about how people would value different services. Taxi service is considered the most convenient service when the cost and time related attributes are kept equal. Therefore, it is assumed that passengers would be willing to pay a few dollars more for the taxi service compared to other services. The constants for shared-taxi and mini-bus services are considered to be the same since the logit model takes into account the additional access and egress time needed for mini-bus. Remember that the main difference between the mini-bus and shared-taxi is that mini-bus service runs between bus-stops rather than the actual origin and destination of the passengers. Given all of the above, the relation between the alternative specific constants for the three services is given as in (31) where it is assumed that passengers are ready to pay \$2 more for taxi. We present two different cases as will be explained in section 5.3 following this assumption.

$$\begin{aligned}
ASC_{\text{shared}} &= ASC_{\text{bus}} \\
ASC_{\text{taxi}} &= \$2 + ASC_{\text{shared}}
\end{aligned}
\tag{31}$$

The value of in-vehicle time parameters, VOT^{IVTT} , are considered to be based on a discrete probability distribution as given in Table 2. The value of out-vehicle time for mini-bus, VOT^{OVT} , is considered to be 1.7 times the VOT^{IVTT} that are listed. The chosen VOT parameters are similar to the willingness to pay figures presented by Kato et al. (2010) who bring together several estimation results in Japan.

The value of late schedule delay is assumed to be higher than the value of early schedule delay and given by $VOT^{\text{SDL}} = \gamma VOT^{\text{SDE}}$ as mentioned in section 4.2.1. It is assumed that $\gamma = 4$ and VOT^{SDE} is assumed to be 20% the VOT^{IVTT} . Therefore, VOT^{SDL} is equivalent to 80% of VOT^{IVTT} . Since FMOD is a reservation-based system, a later departure time means that the passenger will be picked up later from his/her origin. The associated VOT is considered to be less than the one for in-vehicle time since the passenger can be informed about this delay and this time can be spent effectively in other activities.

The distance parameter in the utility of reject option, β_{dist} is assumed to be -0.002 based on the impact of distance on the utility of the taxi service.

Finally, the passengers are assumed to choose one of the offered products or reject to use FMOD based on the choice probabilities provided by the logit model. The resulting choice probabilities constitute a discrete probability distribution and a uniform random number is generated to assign the chosen alternative.

Table 2: Value of in-vehicle time

VOT^{IVTT}	Probability (% in the population)
0.1 \$ /min	30%
0.2 \$ /min	50%
0.3 \$ /min	7%
0.4 \$ /min	7%
0.5 \$ /min	6%

5.3 Experimental results

The experimental setup is designed in order to quantify the performance of the FMOD system. For the FMOD system we consider two versions of the optimization framework. In the first version the model is forced to offer one product for each service, namely the

constraints in (23), or equivalently constraints (28) in the linearized model, are adjusted to be equality constraints. This version will be referred as FMOD-P1 for the presentation of the experimental results. The original version, where the model can choose the size of the menu to be offered, is given under the name of FMOD.

In order to quantify the added value of optimization, we consider a base model which is referred to as NO-OPT. In this version of the model, all feasible products are considered and for each service the product with the highest utility is selected. Therefore, the choice set is not optimized as done with FMOD. One other option would be to consider the full set of feasible products rather than selecting one for each service. However, the logit models for FMOD and NO-OPT would not be comparable in the presence of a different number of alternatives.

Furthermore, we consider scenarios where the fleet has a fixed number of vehicles that can serve each service type. The idea here is to show the added value of the dynamic allocation of the vehicles, i.e. the possibility of changing the role of vehicles during the day. The extreme cases are that all 60 vehicles can only serve taxi, shared-taxi or mini-bus. We further include all the combinations of these services with increments of 10 vehicles. An example would be a fleet of 10 vehicles for taxi, 20 vehicles for shared-taxi and 30 vehicles for mini-bus services that is represented by (10,20,30). We have 28 different cases for such fixed fleet scenarios.

The comparative analysis of FMOD with respect to NO-OPT and fixed fleet scenarios is conducted based on two performance measures: profit and consumer surplus. For consumer surplus we use the logsum of offered alternatives as a measure. The logsum for FMOD is constructed with the optimized list of alternatives and the reject option which is given as follows:

$$\frac{1}{\mu} \ln [\exp (\mu V_{\text{reject}}) + \sum_{n,m,l} x_{n,m,l} \exp (\mu V_{n,m,l})] \quad (32)$$

For the NO-OPT case, we use the alternative with the highest utility for each service to calculate the logsum. Therefore, NO-OPT is equivalent to a model where we maximize the consumer surplus. Note that consumer surplus is in monetary units since the utility functions are normalized as mentioned in section 4.2.1.

In the remainder of this section we provide the comparative analysis for two different scenarios. The first one given in section 5.3.1 assumes relatively lower ASC values for the services. This assumption yields a higher portion of passengers who reject the offered alternatives. The second scenario presented in section 5.3.2 has a relatively lower reject

probability. These two scenarios represent different outcomes for the FMOD system under different assumptions for travel behavior. The analysis of the two scenarios provides interesting insights about the system.

5.3.1 High reject probability

In this scenario, we use the following values for the ASC parameters:

$$ASC_{\text{shared}} = ASC_{\text{bus}} = \$1$$

$$ASC_{\text{taxi}} = \$3$$

The simulation results show that FMOD has 74% higher profit compared to NO-OPT but the consumer surplus reduces by 6%. For FMOD-P1, the profit increase with respect to NO-OPT is 65% since we force the model to offer one product for each service. The consumer surplus reduction compared to NO-OPT is 4.9% which is better than FMOD. This shows the clear trade-off between passenger satisfaction and operator profit.

In Figure 9, we present the proportion of different assortment types offered with the three models we consider. TSB denotes the case where all the three services are included in the assortment and similarly TS has only taxi and shared-taxi. NO-OPT by design has all three services in the assortment when available. FMOD tends to offer taxi and shared-taxi rather than the mini-bus alternative since profit can be increased further without the mini-bus alternative being in the choice set. Remember that taxi has the highest price and mini-bus has the lowest. FMOD-P1 needs to have all the services when available since it is a constraint of the model to have one product per service.

In Figure 10, we present the share of services for each of the models. Shared-taxi has the highest share for all the cases. Another observation is that mini-bus receives the highest share in the case of NO-OPT and the lowest share in the case of FMOD. As mentioned before, the strategy of FMOD is to include fewer mini-bus services in the assortment in order to increase the expected profit. FMOD and FMOD-P1 result in a higher share of taxi compared to NO-OPT, which is again as a result of profit maximization. The share of reject option for NO-OPT is 36% and 38% for both FMOD and FMOD-P1. The number of lost passengers is similar for the three cases. Since the number of lost passengers constitutes a considerable portion of the received requests, the scenario presented in section 5.3.2 investigates the option of having higher ASC values.

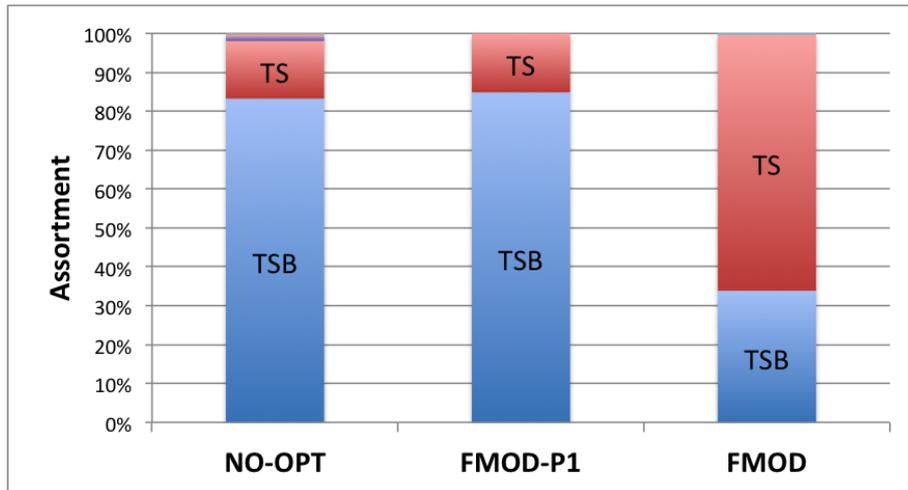


Figure 9: Services included in the assortment with NO-OPT, FMOD-P1 and FMOD

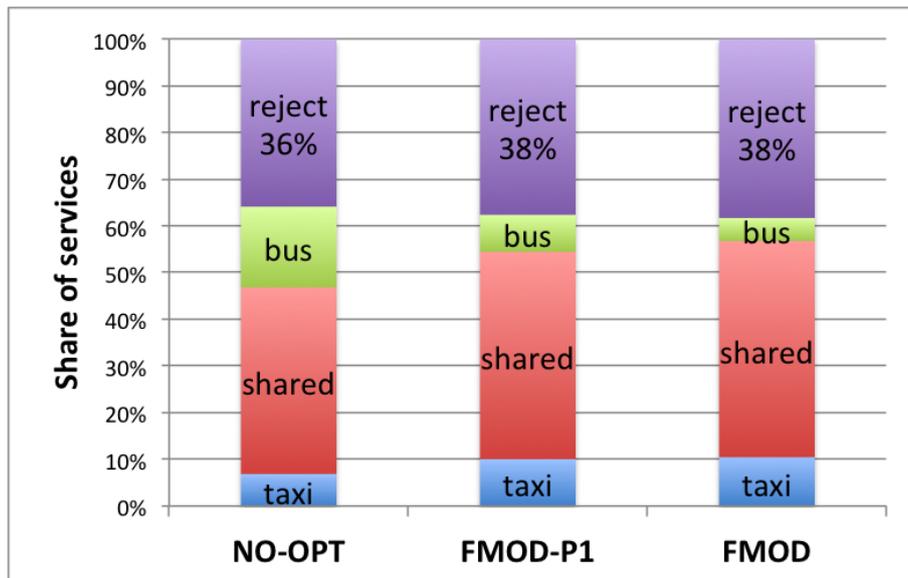


Figure 10: Share of services with NO-OPT, FMOD and FMOD-P1

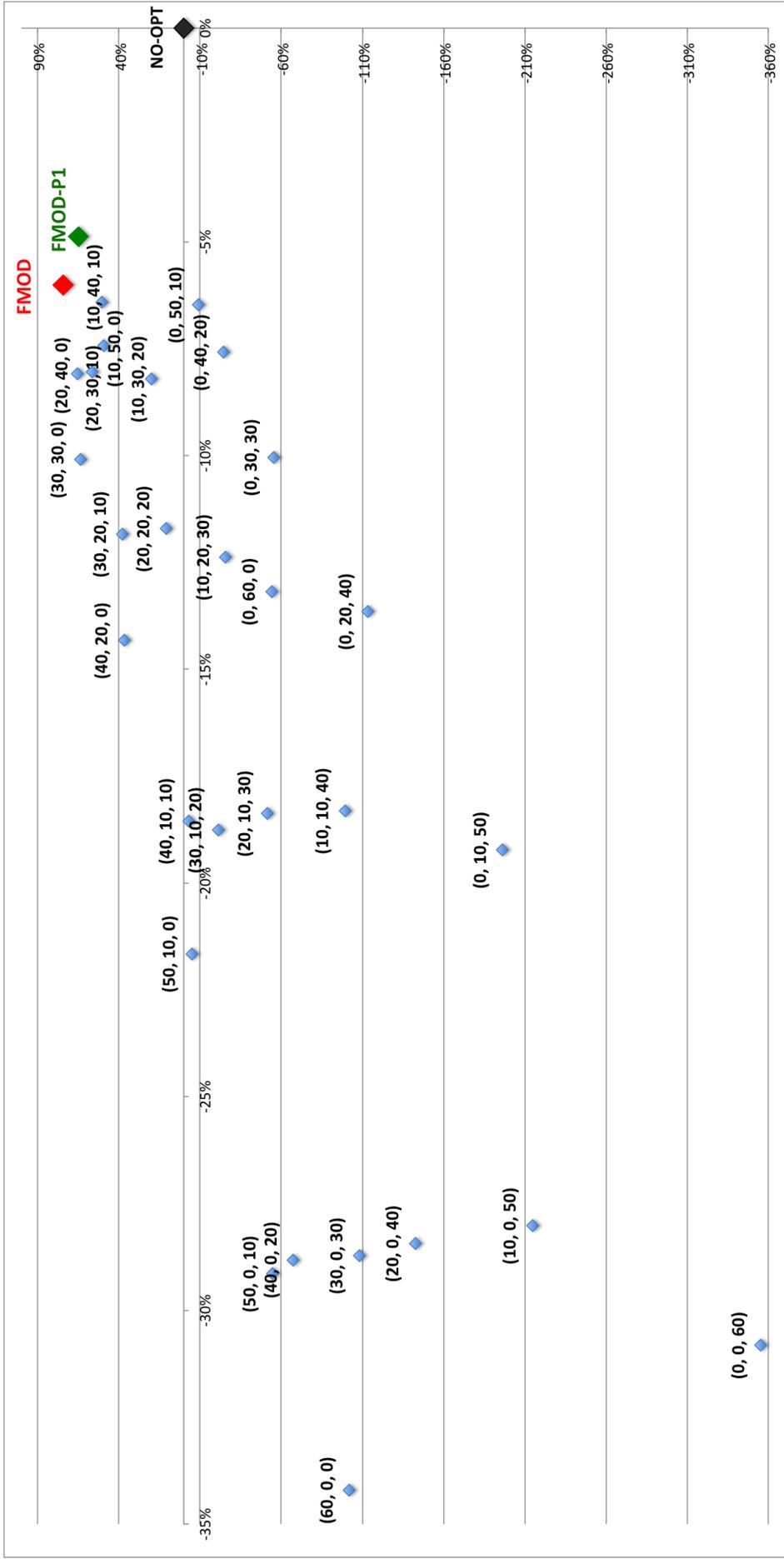


Figure 11: Consumer surplus (x-axis) vs profit (y-axis) with respect to NO-OPT

As mentioned before, we also consider scenarios where the fleet has a fixed number of vehicles that can serve each service type. In Figure 11, we present the results for all the fixed fleet scenarios together with FMOD, FMOD-P1 and NO-OPT. The consumer surplus and profit are presented as percentage differences from the NO-OPT case. NO-OPT has the best consumer surplus since it includes the highest utility products by definition. It is seen that FMOD and FMOD-P1 dominate all the fixed fleet scenarios. All of the fixed fleet scenarios are worse in terms of both profit and consumer surplus. However, there are scenarios such as (20,40,0), (30,30,0), (20,30,10) and (10,40,10) that have closer performance to FMOD and FMOD-P1. These are scenarios where there are more vehicles dedicated to taxi and shared-taxi and few or none serving as mini-bus. Indeed, we can see from the figure that the scenarios with significantly worse profit correspond to the cases where there are many vehicles dedicated to mini-bus service such as (0,0,60), (10,0,50) and (0,10,50). The scenario with the worst consumer surplus corresponds to the case where all the vehicles are dedicated to the taxi service (60,0,0) as expected.

5.3.2 Low reject probability

As mentioned before, the motivation for this scenario with higher ASC values is to analyze the case when the travelers have higher utility towards FMOD services compared to reject option. In the previous scenario, the portion of requests that are lost is more than one third of the received requests. We expect to have lower reject probability with the following values for the ASC parameters:

$$ASC_{\text{shared}} = ASC_{\text{bus}} = \$8$$

$$ASC_{\text{taxi}} = \$10$$

Having higher ASC's for the services allows FMOD to have higher profit by offering assortments that mostly consist of taxi service. Even though this seems better in terms of operator profit, it results in a significantly lower consumer surplus since shared-taxi and mini-bus services are less frequently included in the assortment. In order to have a better passenger satisfaction, a constrained version of the FMOD model is introduced. We define a new constraint for controlling the reject probability based on the reject probability of NO-OPT. We select a threshold value such that the reject probability of FMOD for each request is allowed to be more than the reject probability of NO-OPT up to this threshold value. For example, FMOD-C0 represents the case where the reject probability of FMOD

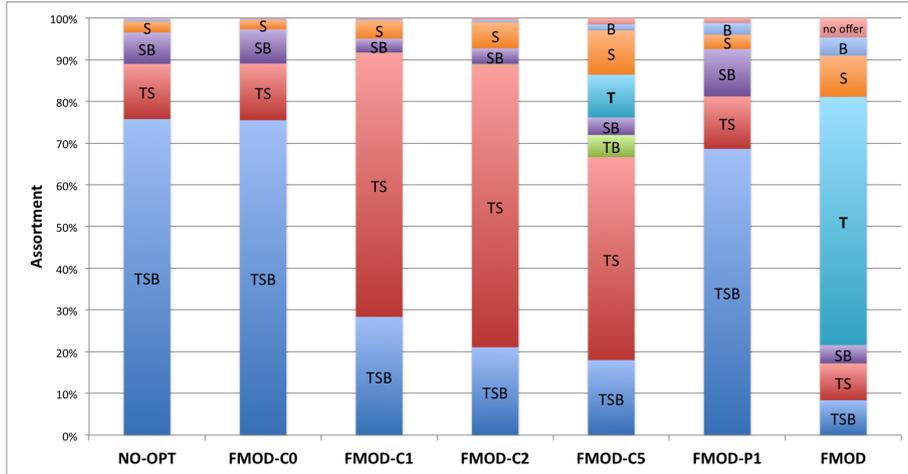


Figure 12: Services included in the assortment

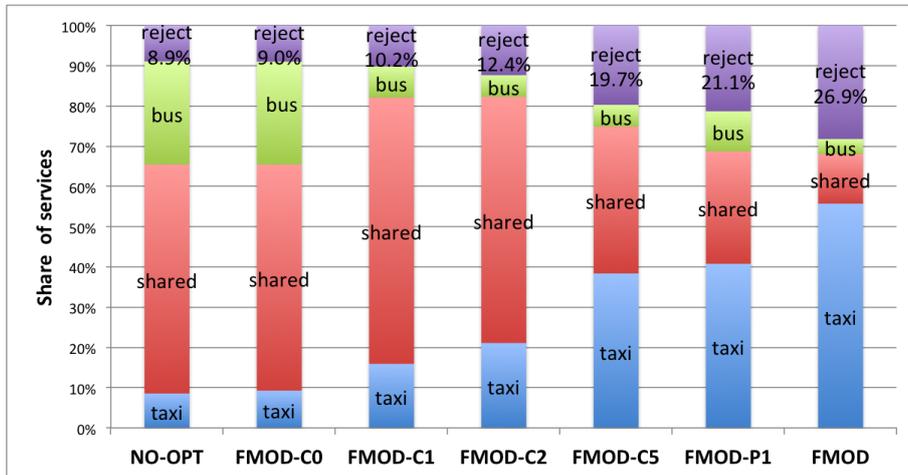


Figure 13: Share of services

cannot be any higher than reject probability of NO-OPT for each ride request and FMOD-C1 is the case where it can be at most 1% higher. Similarly, we introduce FMOD-C2 and FMOD-C5 in the same manner. When there is no feasible assortment under the reject probability constraint, we allow the model to offer the assortment in the original version of FMOD. Note that we did not introduce such versions of FMOD in section 5.3.1 since in the previous scenario the reject probabilities are around the same level for NO-OPT and FMOD and therefore consumer surplus for FMOD is not very low compared to NO-OPT.

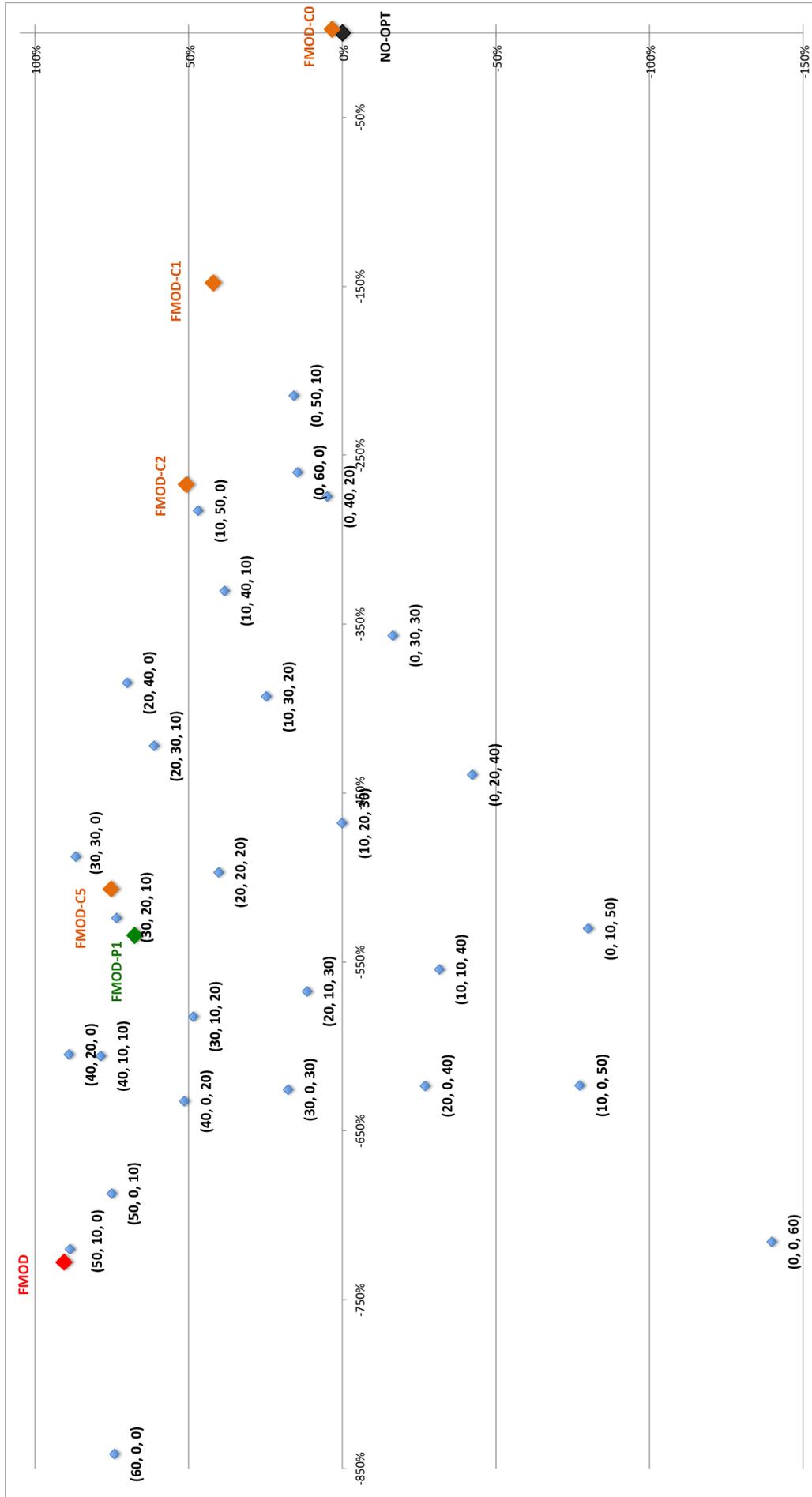


Figure 14: Consumer surplus (x-axis) vs profit (y-axis) with respect to NO-OPT

In Figure 12, we present the proportion of different assortment types for NO-OPT and all the different versions of FMOD. NO-OPT and FMOD-C0 are very close as expected and the assortments mostly include all three services when feasible. FMOD on the other extreme, offers almost 60% of the case taxi-only assortments. When we apply a reject probability constraint, the assortments mostly include taxi and shared-taxi alternatives. Compared to FMOD, these constrained cases can have a higher consumer surplus since shared-ride is more often included in the assortment. FMOD-C5 has taxi-only assortments but this happens to be relatively rare compared to FMOD. This analysis on the offered assortment shows that the importance given to passenger satisfaction considerably alters the strategy of the FMOD system in terms of the offers presented to the passengers.

Figure 13 presents the share of services chosen by the passengers. Compared to the scenario given in section 5.3.1, the reject probability is significantly lower in general. As expected, as we constrain FMOD more, the reject probabilities are closer to the case of NO-OPT. As we move from FMOD-C0 to FMOD-C5 we see that the share of taxi increases and share of shared-taxi and mini-bus decreases. This phenomenon results in a higher reject probability as expected. Note that we apply the reject probability constraint for each ride request with the same value and even though FMOD-C5 has a constraint with a 5 % threshold, the resulting reject probability is around 10.8%. The reason is that in the early time periods FMOD-C5 offers a bit more taxi-only assortments compared to NO-OPT. This results in a capacity shortage later in the day and the reject probability increases in a propagated manner. This phenomenon is also observed in FMOD-C2 but is not as prominent.

Finally, in Figure 14 we present the consumer surplus and profit values for all the models including the fixed fleet scenarios as a percentage difference from NO-OPT. It is observed that FMOD has the highest profit and therefore is not dominated. However, it gives a very poor consumer surplus compared to all others. Constrained FMOD options balance the consumer surplus and profit better so that FMOD-C1 and FMOD-C2 are not dominated by any of the others. NO-OPT has the best consumer surplus by definition and FMOD-C0 is slightly better than that. Similar to section 5.3.1, the lowest consumer surplus occurs when all the vehicles in the fleet are used as taxi (60,0,0) and the lowest profit occurs when all of them serve as mini-bus (0,0,60). There are fixed fleet scenarios which dominate FMOD-C5 and FMOD-P1 such as (40,20,0) and (30,30,0) which have no available mini-bus service. This analysis is important in order to evaluate the trade-off between operator's profit and consumer surplus. In the previous scenario with lower

ASC's, the utility of FMOD services are closer to the reject option and that is why the reject probability is higher. In that case FMOD offers assortments that at least include shared-taxi in addition to the taxi service. However, in this scenario utility towards FMOD services is higher compared to the reject option due to higher ASC's. This motivates FMOD to offer taxi-only assortments which results in lower consumer surplus. This is not preferable since in the long run the operator will face potential passenger losses. Therefore the consumer surplus is controlled through the constraint on the reject probability towards better passenger satisfaction.

6 Conclusions and future research directions

In this paper, we have introduced an innovative on-demand transportation system, FMOD, which provides a menu of travel options to passengers in real-time. The FMOD system integrates scheduling, assortment optimization and choice modeling methodologies in order to optimize the list of travel options offered to each passenger request.

The system is tested through simulation experiments as a proof-of-concept. The performance of FMOD is analyzed in terms of operator profit and consumer surplus. Analyses indicate that profit is significantly increased due to the optimization of offered travel options. The added value of dynamic allocation of vehicles is quantified in comparison to fixed fleet scenarios. FMOD outperforms fixed fleet scenarios in most cases with a better profit and consumer surplus. We find that the flexibility provided by FMOD results in improved operator profit and passenger satisfaction. Therefore, FMOD has great potential to make public transportation more competitive compared to the use of private cars.

In this paper, passenger satisfaction is considered through consumer surplus. The trade-off between consumer surplus and operator's profit is controlled through a constraint on the reject probability. This could also be done through an objective function which has both operator's profit and consumer surplus. We intend to conduct this analysis for FMOD in the near future.

The average processing time for one request is between 1.5-2.5 seconds for the different versions of FMOD presented in this paper. FMOD is designed to run in real-time and the online performance is satisfactory for the considered network. The scalability of the proposed framework should be analyzed in future work in order to assess the real-time performance of the system for different network and fleet sizes.

We believe that the ideas brought about by FMOD can be extended in several ways. First, as mentioned before we work with an optimization model that considers the current

passenger request when optimizing the decisions. An immediate extension that we are working on is the development of models and methodologies that take into account future demand when optimizing the offer to the current request. As an example, if there is a high probability that long distance requests will be received in the near future, it might be better to spare some of the services rather than offering to the current passenger. Since we aim to have a real-time system, this extension should be time efficient and appropriate methodologies should be studied. Such an extension towards a dynamic assortment optimization should carefully take into account the uncertainty in demand in order to provide robust solutions.

The results presented in this paper ignore the traffic conditions on the network. Therefore, another important extension would be the development of models that take into account real-time traffic information so that the FMOD services will be more robust to the changing traffic conditions. Moreover, the demand model can be extended to include additional characteristics of the passengers in order to offer more personalized services. The behavior of passengers can be learned as they use the system multiple times and the demand model can be calibrated similar to the idea of recommender systems.

Acknowledgments

The authors would like to thank Xiang Song (MIT, Sloan School of Management) for his work on the transformation of the assortment optimization model.

References

- G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6):734–749, 2005.
- M. Ben-Akiva, A. De Palma, and P. Kanaroglou. Dynamic model of peak period traffic congestion with elastic arrival rates. *Transportation Science*, 20(3):164–181, 1986.
- C. H. Bhat. Incorporating observed and unobserved heterogeneity in urban work travel mode choice modeling. *Transportation Science*, 34(2):228–238, 2000.
- J. Brake, C. Mulley, J. D. Nelson, and S. Wright. Key lessons learned from recent experiences with flexible transport services. *Transport Policy*, 14(6):458–466, 2007.

- J. F. Cordeau and G. Laporte. A tabu search heuristic for the static multi-vehicle dial-a-ride problem. *Transportation Research Part B: Methodological*, 37(6):579–594, 2003.
- J. F. Cordeau and G. Laporte. The dial-a-ride problem: models and algorithms. *Annals of Operations Research*, 153:29–46, 2007.
- L. Coslovich, R. Pesenti, and W. Ukovich. A two-phase insertion technique of unexpected customers for a dynamic dial-a-ride problem. *European Journal of Operational Research*, 175(3):1605–1615, 2006.
- J. Davis, G. Gallego, and H. Topaloglu. Assortment planning under the multinomial logit model with totally unimodular constraint structures. Working paper, 2013a.
- J. Davis, G. Gallego, and H. Topaloglu. Assortment optimization under variants of the nested logit model. Technical report, Cornell University, School of Operations Research and Information Engineering, 2013b.
- A. De Palma, M. Ben-Akiva, C. Lefevre, and N. Litinas. Stochastic equilibrium model of peak period traffic congestion. *Transportation Science*, 17(4):430–453, 1983.
- R. B. Dial. Autonomous dial-a-ride transit introductory overview. *Transportation Research Part C: Emerging Technologies*, 3(5):261–275, 1995.
- L. Fu. Scheduling dial-a-ride paratransit under time-varying, stochastic congestion. *Transportation Research Part B: Methodological*, 36:485–506, 2002.
- G. Gallego and H. Topaloglu. Constrained assortment optimization for the nested logit model. Technical report, Cornell University, School of Operations Research and Information Engineering, 2013.
- G. Gallego, R. Ratliff, and S. Shebalov. A general attraction model and an efficient formulation for the network revenue management problem. Technical report, Columbia University, New York, NY, 2011.
- H. Kato, M. Masayoshi Tanishita, and Matsuzaki T. Meta-analysis of value of travel time savings: Evidence from japan. In *14th World Conference on Transport Research*, 2010.
- F. S. Koppelman and C. Bhat. A self-instructing course in mode choice modeling: Multinomial and nested logit models. Prepared for U. S. Department of Transportation, Federal Transit Administration, 2006.

- C. Mulley and J. D. Nelson. Flexible transport services: A new market opportunity for public transport. *Research in Transportation Economics*, 25(1):39–45, 2009.
- S. P. Parragh, K. F. Doerner, , and R. F. Hartl. Variable neighborhood search for the dial-a-ride problem. *Computers & Operations Research*, 37(6):1129–1138, 2010.
- P. Rusmevichientong, Z. J. M. Shen, , and D. B. Shmoys. Dynamic assortment optimization with a multinomial logit choice model and capacity constraint. *Operations Research*, 58(6):1666–1680, 2010.
- P. Rusmevichientong, D. B. Shmoys, C. Tong, and H. Topaloglu. Assortment optimization under the multinomial logit model with random choice parameters. forthcoming in *Production and Operations Management*, published online, 2014.
- K. Talluri and G. van Ryzin. Revenue management under a general discrete choice model of consumer behavior. *Management Science*, 50(1):15–33, 2004.
- P. Toth and D. Vigo, editors. *The Vehicle Routing Problem*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2001.
- L. Yang. Modeling preferences for innovative modes and services: A case study in lisbon, 2010.