The concept and impact analysis of a flexible mobility on demand system

Bilge Atasoy a,⇑, Takuro Ikeda b, Xiang Song a, Moshe E. Ben-Akiva a

a Massachusetts Institute of Technology, United States
b Fujitsu Laboratories, Ltd, Japan

A R T I C L E   I N F O
Article history:
Received 4 July 2014
Received in revised form 19 March 2015
Accepted 9 April 2015
Available online 19 May 2015

Keywords:
Demand responsive transit
Flexible transportation services
Personalized transportation services
Shared-taxi
Shared-ride
Mobility on demand

A B S T R A C T
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 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 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. It is shown that the dynamic allocation of the vehicles to different services provides significant benefits over static allocation. Furthermore, it is observed that the trade-off between consumer surplus and operator's profit is critical. The optimization model is adapted in order to take into account this trade-off by controlling the level of passenger satisfaction. It is shown that with such control mechanisms FMOD provides improved results in terms of both profit and consumer surplus.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction and motivation

Flexible forms of transportation are of interest to both passengers and transportation operators for many reasons. A survey in the San Francisco Bay Area revealed that 60% of the respondents were willing to consider personalized demand responsive transit systems as reported by Khattak and Yim (2004). 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 passenger’s 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. Conversely, taxis tend to provide greater flexibility but they are unaffordable on a regular basis for many individuals.

From the perspective of transportation operators, while fixed-route services are cost-effective in urban areas, they become less effective in low-density areas as mentioned by Goodwill and Carapella (2008). Therefore, it is economically
inefficient for transportation operators to provide higher frequency or better coverage in those areas. 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, innovative transportation alternatives with higher profitability and greater flexibility 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 being introduced, a critical issue is how best to ensure the profitability of transport operators while providing a high quality of service.

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. This menu of travel options is an optimized menu which is one distinction of the FMOD system. The flexibility provided to the operators is based on the dynamic allocation of the vehicles to different services depending on the received request. The allocation is such that a vehicle changes its role among different services during the day and this is another distinction of the FMOD system.

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 passenger’s 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 main contribution of this paper is that it provides a proof of concept for the FMOD system using appropriate methodologies. FMOD itself is an innovative concept and the nature of FMOD brings the need for a new modeling framework. The scheduling, routing and assortment optimization models are brought together in order to design such a system which is unique in the context of transportation.

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.
FMOD system 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. An 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. (2013b) 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. (2013a) 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.

3. Concept of FMOD

FMOD is an innovative transportation system that has flexibility in terms of the level of service provided by the vehicles in the fleet as depicted in Fig. 1. The FMOD fleet is assumed to be a homogenous fleet consisting of vans. 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 that are predetermined bus stops 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.

Fig. 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. Fig. 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/destination of the requested trip.
- Preferred departure/arrival time.
- 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 consists of ride alternatives with different service types, depending on the passenger request. The

---

Fig. 2. Components of the FMOD system.
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 is defined by the following attributes:

- Service type (taxi, shared-taxi or mini-bus).
- Pick-up/drop-off location.
- Scheduled pick-up/departure time (DT).
- Scheduled drop-off/arrival time (AT).
- Price.

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 optimized options for his/her requested trip in real-time. 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 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

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

The overall procedure for the feasible product set generation is illustrated by Fig. 4 and the details of the procedure are described together with the constraints in the following sections.

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 \((SB_1, SB_2, \ldots, SB_K)\) each of which consists of a sequence of pick-up and drop-offs. \(K\) denotes the number of \(SBs\) 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. When an \(SB\) is finalized, the vehicle moves from the final destination to the origin of the next \(SB\) under an empty schedule.

Fig. 5(a) shows an example of a sequence of \(SBs\) 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 shorter empty time. Fig. 5(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 \(SBs\) is updated upon a new passenger request in two ways:

Case 1: A new \(SB\) is created and inserted into the sequence of existing \(SBs\). This happens when a new role is assigned to the vehicle. 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 predetermined routes taking into account the accessibility of bus stops. An example is provided in Fig. A.1(a) in Appendix A.

Case 2: A pair of stops (pick-up and drop-off) is inserted into one of the existing \(SBs\). This scenario happens when the new request can be served with one of the already assigned services of the vehicle. For the insertion of new stops on the route, we apply the insertion algorithm that is commonly used for DARP in literature (e.g. Jung et al., 2013) due to its computational efficiency. The insertion algorithm can be applied for shared-taxi. 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. An example is provided in Fig. A.1(b) in Appendix A.
4.1.2. Scheduling and capacity constraints

In this section, we introduce the constraints that need to be satisfied for each feasible product. Fig. 4 illustrates how they are integrated into the feasible product generation.

4.1.2.1. Seat capacity constraints. The number of passengers on board a vehicle cannot be more than the seat capacity of the service assigned to the vehicle. Recall that the fleet consists of homogenous vehicles but the allowed capacity depends on the type of service.

4.1.2.2. Consecutive schedule block constraints. In order to have a feasible product, we must guarantee that there is no conflict in terms of location and time between any set of consecutive schedule blocks, \( SB_k \) and \( SB_{k+1} \). The location of the last stop in \( SB_k \) must be the same location as the first stop in \( SB_{k+1} \). If this constraint 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. Similarly, the arrival time at the last stop of \( SB_k \) should be earlier than the departure time from the first stop of \( SB_{k+1} \).

4.1.2.3. 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 (SDT) and arrival times (SAT) of the product will be referenced as committed departure time (CDT) and committed arrival time (CAT), respectively. The schedule at the time of booking can be modified later based on the requests of other passengers in the case of a shared-ride. However, for the sake of convenience to the passengers, we need to limit this change on the schedule.

For a passenger who specifies a preferred departure time, constraint (2) should be satisfied in order to have a feasible product. Similarly, for a passenger who specifies preferred arrival time, we must satisfy constraint (3). These constraints assume that there is a maximum allowance for deviation denoted by \( T_{\text{max}} \).

\[
\begin{align*}
|\text{CDT}_p - \text{SDT}_p| & \leq T_{\text{max}} \quad \forall p \in F \\
|\text{CAT}_p - \text{SAT}_p| & \leq T_{\text{max}} \quad \forall p \in F
\end{align*}
\]

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. An upper limit on the in-vehicle travel time is introduced as a function of the door-to-door travel time.

4.1.3. Tight/loose feasible products

A product is feasible if the schedule satisfies the above-mentioned constraints 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 \( F_{\text{tight}} \) and the set of loose-feasible products by \( F_{\text{loose}} \) where \( F_{\text{tight}} \cup F_{\text{loose}} = F \).
We denote the departure and arrival time windows by \((PDT^-, PDT^+)\) and \((PAT^-, PAT^+)\), respectively. A tight feasible product must satisfy constraints (4) or (5).

\[
PDT^- \leq SDT_p \leq PDT^+ \quad \forall p \in F^{tight} \\
PAT^- \leq SAT_p \leq PAT^+ \quad \forall p \in F^{tight}
\] (4) (5)

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 the passenger, namely the schedule delay, is defined by \(SD_{\text{max}}\). In other words, the loose-feasible products can be outside the preferred time windows by at most \(SD_{\text{max}}\) time units. Therefore, the departure/arrival time for the feasible products including the loose-feasible ones are as follows:

\[
PDT^- - SD_{\text{max}} \leq SDT_p \leq PDT^+ + SD_{\text{max}} \quad \forall p \in F \\
PAT^- - SD_{\text{max}} \leq SAT_p \leq PAT^+ + SD_{\text{max}} \quad \forall p \in F
\] (6) (7)

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 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

In this section we present the assortment optimization problem where the best assortment of products is selected among the set of feasible products, \(F\), that is generated as described in Section 4.1. 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 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 logit model. 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. As mentioned in Section 3, 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.

\[
V_{n,\text{taxi},l} = \text{ASC}_{\text{taxi}} - \text{price}_{n,\text{taxi},l} - \text{VOT}^{\text{IVTT}} \cdot \text{IVTT}_{n,\text{taxi},l} - \text{VOT}^{\text{SDE}} \cdot \text{SDE}_{n,\text{taxi},l} - \text{VOT}^{\text{SDL}} \cdot \text{SDL}_{n,\text{taxi},l}
\] (8)

\[
V_{n,\text{shared},l} = \text{ASC}_{\text{shared}} - \text{price}_{n,\text{shared},l} - \text{VOT}^{\text{IVTT}} \cdot (\text{IVTT}_{n,\text{shared},l} + \Delta \text{IVTT}_{n,\text{shared},l}) - \text{VOT}^{\text{SDE}} \cdot \text{SDE}_{n,\text{shared},l} - \text{VOT}^{\text{SDL}} \cdot \text{SDL}_{n,\text{shared},l}
\] (9)

\[
V_{n,\text{bus},l} = \text{ASC}_{\text{bus}} - \text{price}_{n,\text{bus},l} - \text{VOT}^{\text{IVTT}} \cdot \text{IVTT}_{n,\text{bus},l} - \text{VOT}^{\text{SDE}} \cdot \text{SDE}_{n,\text{bus},l} - \text{VOT}^{\text{SDL}} \cdot \text{SDL}_{n,\text{bus},l} - \text{VOT}^{\text{OVT}} \cdot \text{OVT}_{n,\text{bus},l}
\] (10)

\[
V_{\text{reject}} = \beta_{\text{dist}} \cdot \text{STD}
\] (11)

- \(\text{ASC}_{\text{taxi}}, \text{ASC}_{\text{shared}}\), and \(\text{ASC}_{\text{bus}}\) are the alternative specific constants for taxi, shared-taxi and mini-bus, respectively. The reject option is considered as the reference and its constant is fixed to zero.
- The utility function is 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, \(\Delta \text{IVTT}\) is defined for this additional travel time. For the mini-bus, the out of vehicle travel 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 $\beta_{\text{dist}}$. 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 $(\text{PDT}^{+}, \text{PDT}^{-})$ and $(\text{PAT}^{+}, \text{PAT}^{-})$, respectively. Therefore, SDE and SDL 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 (12). Similarly, for the requests initiated with an arrival time window, they are given by (13). Fig. 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 $\text{VOI}^{\text{SDL}} = \gamma \text{VOI}^{\text{SDE}}$.

\[
\begin{align*}
\text{SDE}_p &= \max(0, (\text{PDT}^+ - \text{SDT}_p)) \\
\text{SDL}_p &= \max(0, (\text{SDT}_p - \text{PDT}^+)) \quad \forall p \in F \\
\text{SDE}_p &= \max(0, (\text{PAT}^+ - \text{SAT}_p)) \\
\text{SDL}_p &= \max(0, (\text{SAT}_p - \text{PAT}^+)) \quad \forall p \in F
\end{align*}
\]

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

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

where $\mu$ 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:

\[
\begin{align*}
\max & \sum_{n \in N} \sum_{m \in M} \sum_{l \in L} r_{n,m,l} \text{Prob}_{n,m}(x) \\
\text{s.t.} & \sum_{n \in N} \sum_{m \in M} x_{n,m,l} \leq 1 \quad \forall m \in M \\
& x_{n,m,l} \in \{0, 1\} \quad \forall p_{n,m,l} \in F
\end{align*}
\]

The objective function is the expected profit obtained from the list of offered alternatives as given by (15). Constraints (16) 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, the passenger pays the same price for different feasible products of a given service. Therefore, the passenger does not face a trade-off between price and utility of a given service. The trade-off is between different services, namely between level of service and price. For a setting with dynamic pricing this set of constraints would be omitted which is left as future work. Finally, we have the definition of binary $x$ variables in (17) for each product in the feasible set $F$.

![Fig. 6. Disutility of schedule delay.](image)
The presented model in (15)–(17) has a nonlinear objective function and binary decision variables. The nonlinearity is due to the probability term which is a function of the x variables given by (14). As Davis et al. (2013b) presented, this problem can be represented by a linear programming problem with a simple transformation. A new set of decision variables, \( \omega_s \), 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 \( \leq 1 \) for the offered products. \( \omega_{\text{reject}} \) is similarly the probability of rejecting all the products in the assortment. The linear programming formulation is given below:

\[
\begin{align*}
\text{max} & \quad \sum_{m} \sum_{n} \sum_{l} r_{n,m,l} \omega_{n,m,l} \\
\text{s.t.} & \quad \sum_{m} \sum_{n} \sum_{l} \omega_{n,m,l} = 1 - \omega_{\text{reject}} \\
& \quad \sum_{n} \sum_{l} \frac{\exp(\mu V_{n,m,l})}{\exp(\mu V_{\text{reject}})} \leq \frac{\omega_{\text{reject}}}{\exp(\mu V_{\text{reject}})} \quad \forall m \in M \\
& \quad 0 \leq \frac{\omega_{n,m,l}}{\exp(\mu V_{n,m,l})} \leq \frac{\omega_{\text{reject}}}{\exp(\mu V_{\text{reject}})} \quad \forall p_{n,m,l} \in F
\end{align*}
\]

Constraint (19) means that the probabilities sum up to 1 including the reject option. Constraints (20) maintain that only one option is offered for each service analogous to constraints (16) in the original formulation. 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 \( \frac{\omega_{\text{reject}}}{\exp(\mu V_{\text{reject}})} \). This relation is given by constraints (21) for each feasible product. Together with constraint (19), it is maintained that the distribution provided by logit is maintained such that we have choice probabilities relative to the attractiveness of the alternatives and they sum up to 1.

4.2.3. Reference models for the comparative analysis of FMOD

The comparative analysis of FMOD 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 is constructed with the list of alternatives presented to the passenger and the reject option as follows:

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

Note that consumer surplus is in monetary units since the utility functions are normalized as mentioned in Section 4.2.1.

In order to quantify the added value of the FMOD system, we have developed different reference models. The version with the highest utility is selected. Therefore, the assortment optimization model given in (15)–(17), or the linearized version given in (18)–(21), is referred as FMOD.

4.2.3.1. NO-OPT. In order to quantify the added value of optimization, we consider NO-OPT as a base model. In this version, all feasible products are considered and for each service the product with the highest utility is selected. Therefore, the assortment optimization model given in Section 4.2.2 is not utilized for assortment optimization. Indeed, NO-OPT is equivalent to a model where we maximize the consumer surplus.

One other option for NO-OPT 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.

4.2.3.2. FMOD-P1. A second version is considered where the model is forced to offer one product for each service, namely the constraints in (16) are adjusted to be equality constraints. We present the modified assortment optimization model for FMOD-P1 as follows:

\[
\begin{align*}
\text{max} & \quad \sum_{m} \sum_{n} \sum_{l} r_{n,m,l} \text{Prob}_{n,m,l}(x) \\
\text{s.t.} & \quad \sum_{n} \sum_{l} x_{n,m,l} = 1 \quad \forall m \in M \\
& \quad x_{n,m,l} \in \{0, 1\} \quad \forall p_{n,m,l} \in F
\end{align*}
\]

The linearized model given in (18)–(21) can be updated similarly by modifying constraints (20).

4.2.3.3. Static models. In order to evaluate the advantage of dynamic vehicle allocation, i.e. the possibility of changing the role of vehicles during the day, we consider static cases 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 vehicles in the fleet 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 \((10,20,30)\). This setting uses the assortment optimization model given in (15)–(17). However the feasible product set generation phase takes into account the constraints on the total number of available vehicles for each service.

4.2.3.4. Constrained FMOD. Finally, we consider constrained versions of the assortment optimization model in order to have better passenger satisfaction. 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. The constrained version is given as follows:

\[
\begin{align*}
\max & \sum_{n \in N} \sum_{m \in M} \sum_{l \in L} r_{n,m,l} \Prob_{n,m,l}(x) \\
\text{s.t.} & \sum_{n \in N} \sum_{m \in M} \sum_{l \in L} \Prob_{n,m,l}(x) \leq 1 \quad \forall m \in M \\
& 1 - \sum_{n \in N} \sum_{m \in M} \sum_{l \in L} \Prob_{n,m,l}(x) \leq \Prob_{\text{reject}}^{\text{NO-OPT}} + \text{threshold} \\
& x_{n,m,l} \in \{0, 1\} \quad \forall p_{n,m,l} \in F
\end{align*}
\]

where constraint (28) is added to the model. The left hand side represents the reject probability based on the assortment decision and the right hand side has the reject probability of the NO-OPT plus a threshold parameter. The linearized version of the model can be adapted in the same manner. Note that, when there is no feasible assortment under the new constraint, we allow the model to offer the assortment in the initial version of FMOD using the model in (15)–(17).

Different threshold parameters are used for comparison purposes. For example, FMOD-C0 represents the case where the reject probability of FMOD cannot be any higher than reject probability of NO-OPT for each ride request. FMOD-C1 is the case where it can be at most 1% higher. We introduce FMOD-C2 and FMOD-C5 in the same manner.

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 is solved by the LP solver provided in R. The time horizon is considered as 24 h for the simulation.

As a case study we consider the city of Hino in Tokyo which has a land area of approximately 9 km × 8 km with a population of 165,644. The map of the city captured by Open Street Map data is given in Fig. 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

As mentioned before, we 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 that dynamically change their service roles during the simulation. Each vehicle is considered to have 8 seats. However the capacity for the taxi service is 1 passenger only.

The scheduling and assignment of drivers are not considered. It is assumed that there is an available driver for each allocated vehicle. We ignore the traffic conditions on the network so that 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, $T_{\text{max}}$, that is given in constraints (2) and (3) is assumed to be 10 min. This means that the committed departure/arrival times can be changed by at most 10 min due to new arriving passengers who will share the ride. This change in the committed schedule is also limited in terms of the in-vehicle time such that it can be at most twice the time of the taxi service for that particular request.

As mentioned in Section 3, the mini-bus service operates between predefined bus stops rather than serving between the preferred origin and destination locations. We use the actual fixed bus routes in Hino and for bus stops we assume that passengers can board/alight at any intersection. 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, $SD_{\text{max}}$, with respect to the preferred departure/arrival time as given in constraints (6) and (7). $SD_{\text{max}}$ is assumed to be 90 min.

The price for 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.

5.2. Demand model parameters

It is assumed that there are 5000 ride requests in a day which is around 1% of the daily trips in Hino. The time-of-day distribution of the daily demand is given in Fig. 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 min.

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 m 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, and train 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 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

<table>
<thead>
<tr>
<th>Service</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taxi</td>
<td>$5 (base) + $0.5 (per 320 m)</td>
</tr>
<tr>
<td>Shared-taxi</td>
<td>$2.5 (base) + $0.25 (per 320 m)</td>
</tr>
<tr>
<td>Mini-bus</td>
<td>$3</td>
</tr>
</tbody>
</table>
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 (30) 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.

\[
\begin{align*}
\text{ASC}_{\text{shared}} &= \text{ASC}_{\text{bus}} \\
\text{ASC}_{\text{taxi}} &= 2 + \text{ASC}_{\text{shared}}
\end{align*}
\]

(30)

The value of in-vehicle time parameters, $\text{VOT}^{\text{IVTT}}$, are considered to be based on a discrete probability distribution as given in Table 2. The value of out of vehicle travel time for mini-bus, $\text{VOT}^{\text{OVT}}$, is considered to be 1.7 times the $\text{VOT}^{\text{IVTT}}$ that are listed. The chosen $\text{VOT}$ 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 $\text{VOT}^{\text{SDL}} = \gamma \text{VOT}^{\text{SDE}}$ as mentioned in Section 4.2.1. It is assumed that $\gamma = 4$ and $\text{VOT}^{\text{SDE}}$ is assumed to be 20% the $\text{VOT}^{\text{IVTT}}$. Therefore, $\text{VOT}^{\text{SDL}}$ is equivalent to 80% of $\text{VOT}^{\text{IVTT}}$. Since FMO 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 the reject option, $\beta_{\text{dist}}$ 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.

5.3. Experimental results

In 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:

\[
\begin{align*}
&\text{ASC}_{\text{shared}} = \text{ASC}_{\text{bus}} = \$1 \\
&\text{ASC}_{\text{taxi}} = \$3
\end{align*}
\]

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. The profit to operating cost ratios and the total travel time for the three models are also given in Table 3.

In Fig. 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. Similarly, TS has only taxi and shared-taxi and SB has only shared-taxi and mini-bus. 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 Fig. 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.

As mentioned before, we also consider static scenarios where the fleet has a fixed number of vehicles that can serve each service type. In Fig. 11, we present the results for all the static 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 static scenarios. All of the static 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:

<table>
<thead>
<tr>
<th>Model</th>
<th>Profit to operating cost ratio</th>
<th>Total travel time (h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO-OPT</td>
<td>0.24</td>
<td>630.92</td>
</tr>
<tr>
<td>FMOD-P1</td>
<td>0.40</td>
<td>532.40</td>
</tr>
<tr>
<td>FMOD</td>
<td>0.43</td>
<td>520.62</td>
</tr>
</tbody>
</table>
Having higher ASC’s 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. Therefore, we included the constrained FMOD cases in the analysis. The profit to operating cost ratios and the total travel time for all the considered models are given in Table 4. Note that we did not use such constrained versions of FMOD in the previous scenario since 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.

In Fig. 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.

\[
\begin{align*}
\text{ASC}_{\text{shared}} &= \text{ASC}_{\text{bus}} = \$8 \\
\text{ASC}_{\text{taxi}} &= \$10
\end{align*}
\]
Fig. 11. Consumer surplus (x-axis) vs profit (y-axis) with respect to NO-OPT for the high reject probability scenario.

Table 4
Comparison of model outputs for the low reject probability scenario.

<table>
<thead>
<tr>
<th>Model</th>
<th>Profit to operating cost ratio</th>
<th>Total travel time (h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO-OPT</td>
<td>0.62</td>
<td>760.83</td>
</tr>
<tr>
<td>FMOD-C0</td>
<td>0.64</td>
<td>748.48</td>
</tr>
<tr>
<td>FMOD-C1</td>
<td>0.90</td>
<td>700.30</td>
</tr>
<tr>
<td>FMOD-C2</td>
<td>0.94</td>
<td>714.48</td>
</tr>
<tr>
<td>FMOD-C5</td>
<td>1.08</td>
<td>737.73</td>
</tr>
<tr>
<td>FMOD-P1</td>
<td>1.04</td>
<td>718.87</td>
</tr>
<tr>
<td>FMOD</td>
<td>1.18</td>
<td>697.28</td>
</tr>
</tbody>
</table>

Fig. 12. Services included in the assortment for the low reject probability scenario.
Fig. 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 Fig. 14 we present the consumer surplus and profit values for all the models including the static 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. 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 static 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 an optimized menu of travel options to passengers in real-time. The FMOD system integrates scheduling, routing, 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 the static version. FMOD outperforms static 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.

The average processing time for one request is between 1.5 and 2.5 s 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. Furthermore, the extension of the assortment optimization model in order to have a dynamic pricing framework is an interesting future direction. This will enable the system to offer the same service at different prices based on the preferences of the passengers as an improved way of matching supply and demand.

Acknowledgments

The authors would like to thank three anonymous reviewers and the editor for their valuable comments which improved the paper.

Appendix A. Examples for the feasible product set generation phase

Fig. A.1 illustrates the schedule update upon a new ride request from a passenger. In case 1, a new schedule block, SB3, is generated for the vehicle as a taxi service. In case 2, new stops are inserted at locations e and f into the existing schedule block, SB1, under the shared-taxi service. Note that, sj represents the jth stop of SBj.
REFERENCES


Fig. A.1. Demonstration for the update of the schedule blocks and stops.


