



Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Innovative Applications of O.R.

Customer acceptance mechanisms for home deliveries in metropolitan areas



Jan Fabian Ehmke*, Ann Melissa Campbell

Department of Management Sciences, Tippie College of Business, University of Iowa, Iowa City, IA 52242-1000, United States

ARTICLE INFO

Article history:

Received 6 February 2013

Accepted 19 August 2013

Available online 30 August 2013

Keywords:

Routing

Home delivery

Feasibility check

Congestion

City logistics

ABSTRACT

Efficient and reliable home delivery is crucial for the economic success of online retailers. This is especially challenging for attended home deliveries in metropolitan areas where logistics service providers face congested traffic networks and customers expect deliveries in tight delivery time windows. Our goal is to develop and compare strategies that maximize the profits of a logistics service provider by accepting as many delivery requests as possible, while assessing the potential impact of a request on the service quality of a delivery tour. Several acceptance mechanisms are introduced, differing in the amount of travel time information that is considered in the decision of whether a delivery request can be accommodated or not. A real-world inspired simulation framework is used for comparison of acceptance mechanisms with regard to profits and service quality. Computational experiments utilizing this simulation framework investigate the effectiveness of acceptance mechanisms and help identify when more advanced travel time information may be worth the additional data collection and computational efforts.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Online retailing is the fastest growing retail sector in the US, with an overall growth of about 15% in the past year (US Census Bureau, 2012). Sales from online retailers are expected to increase from \$155 billion in 2009 to \$250 billion in 2014. In Western Europe, an even higher yearly compound growth rate of 11% from €68 billion to €114.5 billion over the same years has been predicted (Schonfeld, 2010). Efficient and reliable delivery of orders is crucial for the lasting economic success of online retailers. The particular challenge of planning and executing attended home deliveries, where customers must be home to receive the delivery, has been demonstrated by the failures of Webvan (bankruptcy in 2001) and Publix Direct (shut down in 2003).

Attended home deliveries require customers and logistics service providers to agree on a service time window in order to avoid failure of delivery. For logistics service providers, accepting more deliveries can yield more revenues, but it can make it harder to reliably make deliveries within the service time windows. The same challenge also holds for service technicians, such as plumbers and appliance repairmen, where the attendance of the customer is required in order to conduct a service. We will focus our discussion in this paper on deliveries, but all of the ideas presented here can

be translated to the service technician context, where even more complex constraints such as varying service times and the skill level of technicians must be considered.

Satisfying the tight service time windows for attended home deliveries in metropolitan areas can be especially difficult due to congested traffic networks. Logistics service providers need to consider expected traffic conditions in deciding which deliveries they can reliably complete on time. Many currently use fairly simple ways to account for congestion. They may assign fewer customers per vehicle in high traffic areas, or they may incorporate a standard amount of buffer time before arrivals in order to increase the likelihood of on-time delivery. However, unnecessary buffer times may lead to increased driver and equipment idle time, increased costs, as well as a decreased number of kilometers traveled per hour (Kim, Lewis, & White, 2005). At the same time, too many delivery commitments may be accepted during rush hours, leading to lateness and violation of service time windows. An approach to time window management that incorporates both detailed information on expected congestion and time-dependent vehicle routing has the potential to create better solutions. To the best of our knowledge, this idea has not been explored in the literature and is thus the motivation for our paper.

In this paper, we investigate the interaction of commitment to a service time window and the reliability of actual deliveries. We develop and compare strategies for deciding which requests to accept, building on ideas introduced in Campbell and Savelsbergh (2005). To this end, we introduce several acceptance mechanisms that differ in the amount of travel time information that is incorpo-

* Corresponding author. Present address: Advanced Business Analytics Group, Department of Business Information Systems, Freie Universität Berlin, 14195 Berlin, Germany. Tel.: +49 30 83858731.

E-mail addresses: jan.ehmke@fu-berlin.de (J.F. Ehmke), ann-campbell@uiowa.edu (A.M. Campbell).

rated. We consider both time-dependent and stochastic travel time information in the decision of whether to accept a delivery request. A time-dependent, real-world inspired simulation framework allows for evaluation of the resulting tour plans. The simulation framework enables the comparison of the acceptance mechanisms with regard to maximization of profits and service quality. This enables us to evaluate the extent that time-dependent and stochastic travel time information may be worth the additional data collection and computational efforts. In Section 2, we sketch recent work on the management of service time windows as well as on time-dependent vehicle routing. The Home Delivery Problem in Metropolitan Areas (HDPMA) is presented in Section 3. Acceptance mechanisms for the solution of the HDPMA are introduced in Section 4. Experimental design and computational experiments are discussed in Sections 5 and 6. We present managerial insights and conclude the paper in Section 7.

2. Literature

The fulfillment process for online retailers as described by [Campbell and Savelsbergh \(2005\)](#) is initiated by a customer who places an order and selects a time window for delivery. When the service provider accepts the request, the order is assembled. After a certain “cut-off date”, delivery tours are scheduled. The following overview summarizes literature on the management of service time windows for order placement (Section 2.1) as well as literature on the routing and scheduling of delivery tours (Section 2.2). Order assembly is beyond the scope of this paper and therefore not discussed.

2.1. Management of service time windows

Literature on the management of service time windows is either on a tactical or an operational level. The tactical level is about the design of service time windows, including decisions on the overall number, length and possible overlap as well as the general concept of delivery fees. The design of service time windows aims at the maximization of the captured demand while limiting expected operational costs. [Punakivi and Saranen \(2001\)](#) report that completely flexible, unattended delivery services reduce costs by up to a third relative to attended deliveries with two-hour service time windows. [Campbell and Savelsbergh \(2005\)](#) found that the expansion of 1-hour service time windows to two hours may increase profits by more than 6%. [Agatz, Campbell, Fleischmann, and Savelsbergh \(2011\)](#) introduce an approach that anticipates routing costs in time window design. Given service requirements and average weekly demands for each zip code area of a delivery region, time windows for each zip code area are determined which minimize the expected costs of delivery.

The operational level refers to the actual selection of a service time window by a customer and the commitment of the service provider. Service providers need to decide when to open and when to close service time windows for selection. [Campbell and Savelsbergh \(2005\)](#) introduce several acceptance mechanisms for determining whether a request can be feasibly accommodated in any of the predefined service time windows. Service providers may offer incentives for the selection of certain service time windows that limit expected costs of delivery based on the profitability of a new request with respect to already accepted and still expected requests. [Campbell and Savelsbergh \(2006\)](#) show that incentive schemes can substantially reduce delivery costs, since customers may accept wider service time windows supporting a more efficient combination of requests in vehicle routing and scheduling operations. [Azi, Gendreau, and Potvin \(2012\)](#) examine a sample of future demand scenarios in order to assess the profitability of

dynamically arising requests. The aforementioned papers do not consider time-dependent and stochastic travel time information in the assessment of routing costs and feasibility.

2.2. Routing and scheduling of delivery tours

Once orders have been placed and have a service time window, delivery tours are planned by vehicle routing and scheduling procedures. The corresponding Vehicle Routing Problem with Time Windows (VRPTW) aims at the determination of the optimal tour plan, where (1) every tour departs and terminates at the depot, (2) every customer is visited exactly once by one vehicle, (3) the number of tours is minimal, and (4) every customer is served within its service time window. A secondary objective is to either minimize the total distance traveled or the total travel times of the particular tours. Service time windows can be soft or hard. Soft service time windows may be violated, but violation induces penalty costs in the objective function. Hard service time windows are strictly enforced in order to keep a solution feasible. For the latter, even just finding a feasible schedule with a given number of vehicles is known to be a complex problem ([Hashimoto, Yagiura, & Ibaraki, 2008](#)).

The VRPTW has been extensively studied in the last 25 years. Exact algorithms and model formulations are reviewed by [Baldacci, Mingozzi, and Roberti \(2012\)](#). Metaheuristics provide high-quality solutions with reduced run time; see [Bräysy and Gendreau \(2005b\)](#) for a review paper. [Lau, Sim, and Teo \(2003\)](#) discuss heuristic solution approaches for the VRPTW with a limited number of vehicles. [Pisinger and Ropke \(2007\)](#) elaborate on a general heuristic for vehicle routing problems including the VRPTW. [Pillac, Gendreau, Guéret, and Medaglia \(2013\)](#) provide a review on dynamic variants of the vehicle routing problem, distinguishing formulations and solution approaches based on the degree of dynamism of considered information.

For vehicle routing and scheduling in metropolitan areas, literature on the time-dependent extension of the VRPTW is relevant. The TDVRPTW captures costs and travel times between depot and customer locations based on the time of the day. Thanks to the availability of more accurate travel time information, there has been an increasing number of papers in the last 10 years that consider information on congestion, e.g., [Ando and Taniguchi \(2006\)](#), [Donati, Montemanni, Casagrande, Rizzoli, and Gambardella \(2008\)](#), [Fleischmann, Gietz, and Gnatzmann \(2004\)](#), [Haghani and Jung \(2005\)](#), [Hashimoto et al. \(2008\)](#), and [Ichoua, Gendreau, and Potvin \(2003\)](#). [Maden, Eglese, and Black \(2010\)](#) present a metaheuristic solution approach (“LANTIME”). [Ehmke, Steinert, and Mattfeld \(2012a\)](#) adapt the LANTIME heuristic for usage with taxi speed data for time-dependent routing in metropolitan areas. [Kok, Hans, and Schutten \(2012\)](#) investigate the impact of time-dependent travel times on the reliability of tour plans. Their experiments underline the necessity of time-dependent vehicle routing in metropolitan areas in order to provide efficient and reliable delivery tours.

3. Home delivery problem in metropolitan areas

In this section, we define the integrated routing and scheduling problem that is studied in this paper. The idea is to maximize the number of requests accepted for delivery by enhanced feasibility checks that provide immediate feedback to a customer during the order process. In particular, we investigate how different levels of travel time information may improve the evaluation of whether a new delivery request is feasible in a potential tour plan. From a fulfillment process perspective, this corresponds to the operational management of service time windows.

Let T be a set of *time slots* the logistics service provider offers for delivery on a particular day. Time slots and their lengths have been predefined through tactical management of service time windows. They are non-overlapping and cover an entire driver's shift, e.g., 12:00–12:30, 12:30–13:00, ..., 19:30–20:00. Let C be a set of customers that demand delivery on a day. Each customer i selects a *customer time window* w_i from T , generating a request r_i with $[early_i, late_i]$ denoting the earliest possible start of service and $late_i$ the latest possible start of service. Requests arrive in real time and are considered up to a certain cut-off time, which precedes the actual execution of delivery tours. Each request r_i is *accepted* or *rejected* the moment it arrives in order to enable immediate customer feedback. If accepted, r_i would be included in the set of accepted requests A . The logistics service provider is then committed to the request, and the request cannot be canceled in order to accept another request instead. In practice, a service provider would offer an alternative time slot if a request for the selected time slot was rejected. We imitate that by allowing the customer to define an alternative request r_i^{alt} for the same day with an alternative customer time window w_i^{alt} and corresponding earliest and latest start of service $[early_i^{alt}, late_i^{alt}]$, which will be checked for feasibility if r_i is rejected. If one of these windows is feasible, w_i or w_i^{alt} will transform into a hard constraint for the subsequent vehicle routing and scheduling problem and will thus play an important role for the feasibility and efficiency of actual delivery tours.

The HDPMA aims at the maximization of the number of accepted requests $|A|$ while considering feasibility of tour plans. Let \mathcal{O} define an oracle that evaluates the feasibility of each request r_i . Whenever a new request r_{new} is posed, \mathcal{O} determines the feasibility of r_{new} based on the set $A \cup \{r_{new}\}$. We investigate static and dynamic manifestations of \mathcal{O} that differ in the level of travel time information considered. While static approaches refer to a rough estimate on feasibility, dynamic approaches consider an actual solution of the TDVRPTW in the decision whether r_{new} should be accepted or not. For the solution of the TDVRPTW, we assume that capacity constraints are not binding. For attended home deliveries, service technicians, and logistics service providers delivering standardized, small-sized parcels, time constraints are typically more limiting than vehicle capacity. Removing capacity restrictions also allows us to focus on the sole impact of more detailed travel time information on advanced acceptance mechanisms.

4. Solving the home delivery problem in metropolitan areas

The decision on whether to accept a request or not is made by an *acceptance mechanism*. We present two different types of acceptance mechanisms to solve the HDPMA. Both provide a quick decision on whether a new request may be accommodated based on the already accepted set of requests. They differ with respect to the extent that travel time and routing information is considered. Static acceptance mechanisms refer to a very rough estimation of expected routing efforts. They imitate simple policies that are often used in practice in order to keep the number of accepted requests in accordance with logistics constraints. Dynamic acceptance mechanisms consider either time-dependent or both time-dependent and stochastic travel time information. Based on expected travel times, they involve finding an actual solution of the TDVRPTW in order to evaluate the feasibility of a request.

4.1. Static approaches

For static approaches, the general idea is to accept a fixed limit of potentially feasible requests max for each time slot t . This value max expresses a conservative estimate of what the service provider believes they could manage in the most resource demanding time slot of the day, considering best practices and constraints such as

the length of time slots, the average duration of service, the maximum number of service vehicles, and the size of the service area. The *SLOT* mechanism would accept a request for a time slot t whenever the number of already accepted requests for this time slot $|A_t|$ plus r_{new} is below or the same as max (as in Campbell and Savelsbergh (2005)). If w_{new} cannot be accommodated, w_{new}^{alt} is checked. If all potential time slots are “full”, r_{new} is rejected.

To account for time-varying congestion in metropolitan areas, we also let the value max vary for different time slots t , since a logistics service provider may likely be able to manage a larger number of requests in non-peak hours than in rush-hour conditions (*TD-SLOT*). To this end, we allow the number of potentially feasible requests to be time-dependent, i.e., max_t varies depending on the expected evolution of travel times for particular time slots. A congested road network can be represented by a smaller max_t for time slots close to rush hours, whereas in non-peak conditions, a higher max_t may reflect increased service capabilities.

TD-SLOT requires general knowledge on the evolution of travel times in a metropolitan area. This information may be provided by a local traffic control center or by digital roadmap vendors like TomTom, who offer time-dependent travel times (“speed profiles”) for major roads in 45 countries (TomTom, 2012). While max and max_t are set in a conservative way for *SLOT* and *TD-SLOT*, note that both still do not guarantee the provision of a feasible tour plan, since there is no control in the assignment of customer requests beyond the number of already accepted requests. Thus, it is possible that a tour plan for a particular set of customer requests turns out to be infeasible in practice. On the other hand, actual utilization of service vehicles may be far from being efficient, since max and max_t reflect only rough estimates of service capabilities, which might lead to preventive rejection of actually feasible requests.

4.2. Dynamic approaches

Dynamic approaches alleviate the disadvantages of static approaches through dynamic feasibility checks. Based on the combined set $A \cup \{r_{new}\}$, an actual TDVRPTW solution is computed and checked for feasibility. If there is a feasible solution, r_{new} is accepted; otherwise, it is rejected. Note that unlike static approaches, detailed information on expected travel and arrival times is exploited. Due to the large amount of additional information that is considered, one would expect a better utilization of logistics resources, while service quality and feasibility are maintained.

To conduct dynamic feasibility checks, we implement a time-dependent variant of the well-known I1 insertion heuristic (Solomon, 1987), which is at the heart of many complex vehicle routing heuristics (Bräysy & Gendreau, 2005a). We chose this heuristic because it was shown to work well in a time-dependent setting (Ehmke et al., 2012a). We note that any off-the-shelf routing solver could be embedded if modified to incorporate dynamic feasibility checks. The following three variants of dynamic acceptance mechanisms are applied:

- *DYN* considers the combined set $A \cup \{r_{new}\}$ in the decision on whether to accept r_{new} or not. The decision is made based on expected, time-dependent travel times and the resulting feasibility of the TDVRPTW solution provided by I1. This is a significant extension of the *DYN* variant proposed in Campbell and Savelsbergh (2005). If the number of required vehicles exceeds the maximum number of vehicles, r_{new} is rejected. Note that this dynamic feasibility check may lead to a very efficient utilization of logistics resources and dense tour plans, contrasting potential underutilization occurring in static approaches.
- In practice, travel times are subject to a variety of environmental and traffic flow inherent influences that may not be captured by a deterministic TDVRPTW solution. To reduce the likelihood

of lateness that might arise in tour execution, we introduce DYN-SBF which extends DYN so that it includes a fixed buffer time bt at each customer i . Considering the scheduled arrival time arr_i , DYN-SBF ensures that $arr_i + bt \leq late_i$ for any request i . Thus, scheduled start of service for each corresponding customer i is at least bt minutes before its latest start of service. Note that the use of a buffer in DYN-SBF is not dependent on the location of the request or the position of the request in a tour. The same buffer is used for all requests.

- **DYN-BUF** aims at alleviating lateness in tour plans by considering both time-dependent and stochastic travel time information. The idea is to reject requests that might lead to a tour plan being prone to lateness and thus provide poor customer service. For instance, DYN-BUF would avoid dense tour plans where scheduled arrival times are close to the end of service time windows, especially after a long distance was driven or when customers are serviced at the end of a long tour. Details of this enhanced feasibility check are given in Section 4.3. In addition to previous approaches, DYN-BUF requires also information on the variation of time-dependent travel times. These can be derived from GPS equipped service vehicles, for example. Corresponding data collection and processing technology is described in Ehmke, Meisel, and Mattfeld (2010).

4.3. Avoiding lateness

DYN-BUF aims at avoiding lateness in the execution of delivery tours by considering time-dependent travel times and their variation in the maximization of the number of requests. For each customer i , an individual buffer time bt_i is computed, contrasting the fixed buffer time bt incorporated by DYN-SBF. The use of bt_i accounts for the impact of travel time variation on the feasibility of visiting customer i within its service time window. It also includes information on travel time variation due to previous customers on the tour. The latter is referred to as *lateness propagation*. The idea is that in a tour with no or little idle time, travel time variation cumulates in the course of a tour, affecting the variation of arrival times at later customers. In a tour featuring a sufficient proportion of idle time, though, the variation of travel times in an earlier part of the tour might have a negligible impact on arrival times at subsequent parts of the tour.

The computation of bt_i requires (1) a way to propagate arrival time variation in the course of a tour and (2) the modeling of the impact of arrival time variation on the feasibility of visiting a customer within its service time window. Chang, Nozick, and Turnquist (2005) and Chang, Wan, and OOI (2009) approximate the distribution of arrival times in the course of a tour by adding means and variations of travel times according to the convolution of the normal distribution. They highlight that this approach provides a sufficient approximation of arrival time variations, even if travel time distributions for individual paths do not follow a particular type of distribution. Following this idea, we can formulate the *cumulated standard deviation of arrival time* $\hat{\gamma}_i$ at customer i as

$$\hat{\gamma}_i = \sqrt{\hat{\gamma}_{i-1}^2 + \sigma_{i-1,i}^2}, \hat{\gamma}_0 = 0,$$

with $\sigma_{i-1,i}^2$ reflecting the standard deviation of travel times for the path from the previous customer $i-1$ to the current customer i . The value $\hat{\gamma}_{i-1}$ denotes the cumulated standard deviation that arises from more previous paths (from the depot to $i-1$). Whenever a request is checked for feasibility, the buffered arrival time $\hat{b}t_i$ can then be computed as $\hat{b}t_i = \alpha \hat{\gamma}_i$ with α as a user specific service level reflecting the extent that the variation of travel times is considered. Note that this simple computation of $\hat{b}t_i$ does not consider the proximity of the mean arrival time arr_i to $early_i$ or if idle time will be likely at i . This is included as follows.

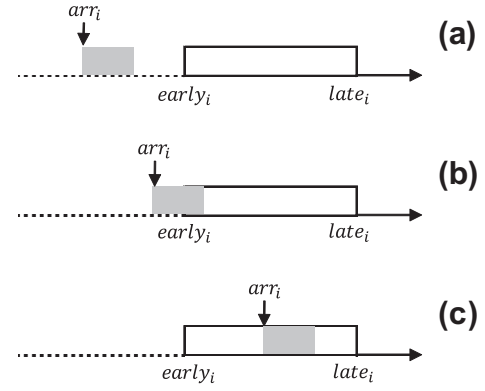


Fig. 1. Arrival time variation and service time windows.

A TDVRPTW solution provides information on the mean arrival time arr_i for each customer i . Let β_i denote the proportion of the distribution of arrival times around arr_i that overlaps with the service time window of customer i :

$$\beta_i = 1 - \Phi(z) \quad \text{with} \quad z = |arr_i - early_i| / \hat{\gamma}_i$$

If arr_i is significantly before $early_i$, we assume that there is enough idle time to compensate for cumulated variation of arrival times (cf. Fig. 1(a)). Then, the cumulated variation of arrival times should not be propagated, which is reflected through $\beta_i = 0$. If arr_i is closer to $early_i$, a proportion of the distribution of arrival times arr_i overlaps with the service time window, being reflected by $\beta_i \in (0; 1)$ (cf. Fig. 1(b)). If arr_i is within the service time window, $\beta_i = 1$ ensures that the cumulated variation of arrival times is fully propagated.

Considering the cumulated standard deviation of arrival time and its impact on the feasibility of visiting customer i in its service time window, bt_i can be written as

$$bt_i = \alpha \beta_i \gamma_i \quad \text{with} \quad \gamma_i = \sqrt{\beta_{i-1} \gamma_{i-1}^2 + \sigma_{i-1,i}^2} \quad \text{and} \quad \gamma_0, \beta_0 = 0.$$

The multiplier β_{i-1} accounts for the proximity of arr_{i-1} to $early_{i-1}$ and adjusts the cumulated standard deviation γ_{i-1}^2 accordingly. If $arr_i + bt_i > late_i$ for some customer i within a potential tour plan, the tour plan would be handled as infeasible, and r_{new} would be rejected.

5. Experimental design

Evaluation of acceptance mechanisms is conducted through computational experiments. We first explain the construction of problem instances and how they evolve into actual tour plans (Section 5.1). For vehicle routing, we introduce a road network model based on time-dependent travel times. Then, the evaluation of tour plans by simulation is described (Section 5.2). For simulation, time-dependent travel time distributions are presented. We finally list the metrics for the evaluation of tour plans (Section 5.3).

5.1. Construction of problem instances

To derive expected travel times for vehicle routing, a real world inspired road network model is developed based on large amounts of travel time data from the metropolitan area of Stuttgart, Germany. Customer locations are represented by nodes of a digital roadmap. They are distinguished by two different zones, namely the *downtown* and the *suburban area*. We thus capture typical variation in the proximity of customers, which is much closer in downtown areas than in suburban areas. Furthermore, different travel time patterns are taken into account, e.g., congestion from suburban areas in the direction of downtown areas in morning

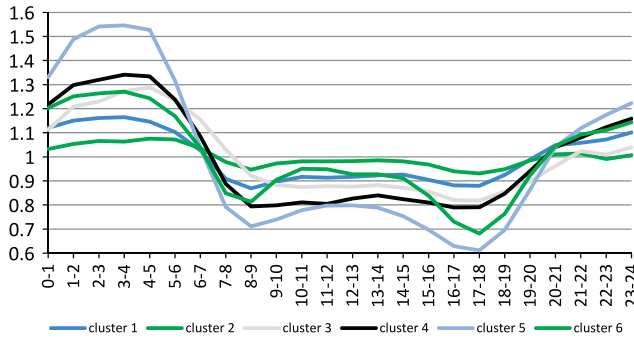


Fig. 2. Speed multipliers for a metropolitan area (Ehmke et al., 2012b).

peak hours. Straight-line distances between nodes are multiplied with a correction factor of 1.5, which is a rough approximation of a typical shortest path distance on the underlying road network. These distances are then converted into travel times based on typical speeds within and between zones. Travel times are adjusted by time-dependent multipliers for a typical weekday in order to capture time-varying congestion (cf. Fig. 2). Speeds and multipliers are derived from analysis of large travel time databases conducted by Ehmke, Meisel, and Mattfeld (2012b). Multipliers for the adjustment of downtown paths are provided by “cluster 6”, while multipliers for the adjustment of suburban paths are provided by “cluster 5”. The corresponding clusters mainly contain inner city and main urban roads, respectively, and serve as a simplified representation of a complex road network of a metropolitan area. Different multipliers reflect the typically varying extent of speed variation at different times of the day.

Construction and processing of requests work as follows. For each zone of customer locations (downtown/suburban), n customers are randomly selected. The order of selection defines the order of arrival of new requests for processing in acceptance mechanisms. Two different time window options w_i and w_i^{pl} are then randomly assigned to each customer from the fixed set of time slots T , transforming a customer location into a request. For static approaches, a tour plan is not created until A is final. For dynamic approaches, a tour plan is created whenever a new request is posed, i.e., for each combined set $A \cup \{r_{new}\}$. Resulting tour plans provide information on the number of required vehicles, total distance and travel time, as well as scheduled departure and arrival times at customers.

5.2. Tour plan simulation

Although we consider time-dependent travel times in the generation of tour plans, actual feasibility and service quality might suffer from travel time variation arising in practice. To evaluate generated tour plans with respect to service quality and feasibility in real-life conditions, we use a time-dependent and stochastic simulation framework. Simulation of tour plans reveals the average ratio of tour plans where lateness occurs, the average number of time window violations, the average amount of lateness when a time window violation occurs, as well as the worst-case lateness.

For our tour plan simulation, we need to model urban travel times in a reasonable way. Following Susilawati, Taylor, and Somehahalli (2011), who conduct a comprehensive statistical analysis of travel time variability on urban arterial roads, it is assumed that travel times of paths in metropolitan areas follow the Burr XII Type 3 distribution. The cdf of Burr XII Type 3 is given as

$$F(x) = 1 - \left(1 + \left(\frac{x}{k}\right)^j\right)^{-l}.$$

The inverse function of the Burr XII Type 3 cdf can be computed as

$$F^{-1}(x) = k((1-x)^{-1/l} - 1)^{1/j}.$$

We were able to estimate pdfs and cdfs for non-peak and peak hours based on standardized path durations arising from approximately 23,000 path realizations provided by Ehmke et al. (2012b). For nearly all sets of paths, statistical analysis reveals the superiority of Burr XII Type 3 compared with well-known travel time distributions such as Normal, Lognormal, Gamma or Weibull. Two examples of fitted pdfs for suburban peak and non-peak times are shown in Fig. 3. The x-axis presents the standardized duration of a path. Note the relatively long tail of the pdfs, which is typical of Burr XII Type 3 distributions.

The actual simulation of tour plans follows the parameters given in Table 1. We generate a time-dependent travel time for each path of a tour by evaluation of $F^{-1}(x)$ with x being randomly selected from the interval $[0; 1]$. The distribution of travel times for each path depends on its departure time and the location of origin and destination. Distributions for peak hours are used for departure times between 7:00 and 9:00 as well as between 16:00 and 18:00. In case of late arrival at a customer i , we note the occurrence as well as the amount of lateness. The amount of lateness is given by the difference between simulated arrival time and the latest start of service $late_i$.

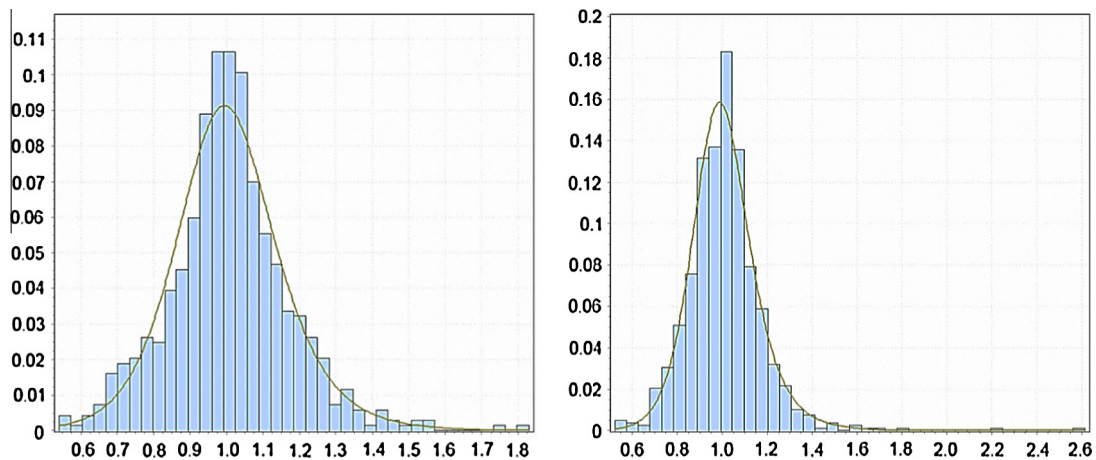


Fig. 3. Fitted pdfs for standardized travel times of suburban-downtown paths (peak/non-peak pdf).

Table 1
Parameters for simulation of paths according to Burr XII Type 3 distribution fit.

Time period	j	k	l
<i>Downtown → downtown</i>			
Off-peak	30.348	0.9810	0.4404
Peak	14.408	1.0105	0.8711
<i>Suburban → downtown</i>			
Off-peak	11.666	1.0109	1.0980
Peak	9.842	0.9538	0.6782
<i>Suburban → suburban</i>			
Off-peak	23.795	0.9877	0.4799
Peak	10.330	0.9836	0.6235
<i>Downtown → suburban</i>			
Off-peak	9.842	0.9538	0.6782
Peak	11.666	1.0109	1.0980

5.3. Metrics

Effectiveness of acceptance mechanisms is evaluated through analysis of a variety of problem instances. For each problem instance, we report on (1) the characteristics defining the instance, (2) results arising from accepting and rejecting customers in terms of tour plans, and (3) results from the simulation of tour plans.

5.3.1. Instance characteristics

For SLOT and TD-SLOT, the *number of potentially feasible requests* is reported, which reflects a rough estimate on the number of requests a logistics service provider can accommodate within a business day. For SLOT, the *number of potentially feasible requests* is $\max * \text{number of service time windows}$. For TD-SLOT, it is the sum of the \max_t values. The values of \max and \max_t are determined by incrementally increasing them until resulting tour plans tend to become infeasible. Note that there is not such a predefined number for dynamic solution approaches. The *number of posed requests* denotes the number of customer requests that are subject to acceptance/rejection decisions. Each request comes with a fixed *service time* and a fixed *length of the service time window*, which corresponds to the length of the underlying time slot. Furthermore, a request is classified according to its *customer location* (downtown/suburban).

5.3.2. Customer acceptance

LDPMA solutions are presented as follows. We report on the *number of accepted requests* as well as on the number of accepted requests whose *first time window option* is accepted. To make results more comparable, we denote the ratio of these numbers with respect to the maximum number of accepted requests of all acceptance mechanisms, which is provided by DYN. We also characterize resulting tour plans by their *total distance* and *denseness*. Denseness is the ratio of the sum of travel and service times to the total duration of the tour plan. A high value of denseness represents a very efficient tour plan with limited idle time, whereas a low value of denseness would represent a rather inefficient tour plan including a large proportion of idle time. We furthermore report on the *average arrival of rejected requests* relative to the first and last posed request (on a relative scale between 0 and 1). Finally, the *average distance between rejected customers* and the depot is provided. These last two values help characterize rejected requests in a temporal and spatial manner.

5.3.3. Simulation of tour plans

The simulation of tour plans provides information about the lateness that might arise in real-life conditions. We report on the *proportion of simulations* where a tour plan runs late, i.e., where lateness occurs at any customer. Furthermore, the *average number*

of *time window violations* per simulation run is given. We also report on the *average lateness at customers* where lateness occurs as a measure of service quality. The average value of the *95th lateness percentile* characterizes worst-case situations in the execution of delivery tours.

6. Computational experiments

Based on the experimental design introduced in Section 5, we investigate the effectiveness of acceptance mechanisms by presenting metrics on resulting tour plans as well as on the simulation of tour plans. Instances are analyzed with respect to varying locations of customers (Section 6.1), varying lengths of customer service times (Section 6.2), varying lengths of service time windows (Section 6.3), the impact of lateness avoidance (Section 6.4), as well as varying customer behavior (Section 6.5).

For each experiment, the average results of 100 different sets of requests are presented. For static acceptance mechanisms, average results stem from tour plans that have been computed by I1 after completion of the corresponding acceptance process. Each of the 100 resulting tour plans is simulated 1000 times. To ensure comparability of results, the number of posed requests is always 25% larger than the total number of potentially feasible requests predefined for SLOT. The size of the service fleet is set to three vehicles. DYN-SBF is parameterized with $bt = 5$ because it is a value that works sufficiently well for the majority of instances. To establish a moderate level of service, we set $\alpha = 2.0$ for DYN-BUF.

6.1. Varying customer locations

The locations of customers might have a significant impact on the efficiency and reliability of tour plans. In Table 2, average results from 100 downtown instances are compared to 100 suburban instances. For both types of instances, the service time per request is fixed to 20 minutes. The length of a service time window is 60 minutes, reflecting eight non-overlapping service time slots between 12:00 and 20:00. Each acceptance mechanism processes 70 requests for downtown instances and 50 requests for suburban instances in total. SLOT would accept up to $\max = 7$ requests per time slot in downtown areas and up to $\max = 5$ requests in suburban areas, i.e., up to 56/40 requests in total, reflecting the larger spread of customer locations in suburban areas. TD-SLOT would accept up to 58/42 requests in total, varying \max_t according to the expected evolution of travel times in the course of the day. The value for \max and the values for \max_t were determined by preliminary investigation such that a feasible tour plan can likely be provided. For both SLOT and TD-SLOT, a larger number of accepted requests would often lead to infeasible tour plans.

On average, the different acceptance mechanisms accommodate between 55.6 and 62.2 requests for downtown customer locations and between 39.4 and 47.2 requests for suburban customer locations, reflecting the difference in distances between customers. TD-SLOT is able to accept slightly more requests than SLOT. A significant increase in the number of accepted requests is provided by DYN and DYN-BUF.

For downtown customer locations, the denseness of DYN generated tour plans is relatively high (96%), accompanied by a relatively large proportion of tour plans running late (56%). DYN-BUF is able to reduce this value to 39% with a very small reduction in the number of accepted requests. In general, the extent of lateness is fairly negligible in downtown areas. The average number of time window violations is below one and the average lateness is nearly zero. This is due to close customer locations and a relatively low speed level. DYN-SBF yields a significantly smaller number of requests (91% of DYN) while eliminating late tour plans. On average, it re-

Table 2
Results for downtown and suburban instances.

Customer location	Acceptance mechanism	# Of feasible requests	# Of posed requests	Service time	Length of time window	Avg # accepted requests	Avg # requests rel. to DYN (%)	Avg # of first options	Avg # of first options rel. to DYN (%)	Tour plan distance	Tour plan denseness (%)	Avg arrival of rej. requests	Avg dist. of rej. customers	Proportion tour plans late (%)	Avg # of tw violations	Avg late at customer	Avg 95th late at customer
Downtown	SLOT	56	70	20	60	55.6	90	48.3	78	110	88	0.83	4.1	9	0.1	0.0	0.8
	TD-SLOT	58				57.5	92	49.7	80	113	90	0.83	4.1	22	0.3	0.0	1.4
	DYN-SBF	–				56.5	91	45.5	73	108	89	0.74	4.3	0	0.0	0.0	0.4
	DYN-BUF(2.0)	–				62.0	100	55.4	89	114	95	0.85	4.2	39	0.5	0.0	2.2
	DYN	–				62.2	100	56.6	91	114	96	0.88	4.2	56	0.8	0.0	2.5
Suburban	SLOT	40	50	20	60	39.4	83	33.7	71	284	81	0.81	8.8	59	0.9	4.1	12.0
	TD-SLOT	42				40.8	86	34.7	73	289	82	0.81	8.9	55	0.8	3.3	9.5
	DYN-SBF	–				42.1	89	33.9	72	291	84	0.71	10.2	35	0.6	4.9	14.7
	DYN-BUF(2.0)	–				45.6	97	39.1	83	304	89	0.75	10.3	70	1.4	4.1	12.4
	DYN	–				47.2	100	44.0	93	310	91	0.90	9.5	86	2.3	4.0	12.3

Non-overlapping delivery time windows for a typical weekday, 12 pm–8 pm; fleet size = 3 vehicles.

jects requests relatively early (at point 0.74). Rejected requests are about the same distance from the depot (4.1–4.3 kilometers) for all mechanisms.

For suburban instances, distances of tour plans nearly triple. Denseness of tour plans decreases to a maximum of 91% (DYN), reflecting challenges that arise from the large spread of customer locations in suburban areas. Static approaches perform relatively poorer than for downtown instances with accepting only 83/86% of the number DYN would accept. Due to longer distances traveled and a higher impact of congestion, lateness occurs much more often than in downtown areas. For DYN, 86% of simulated tour plans run late, resulting in 2.3 violations of time windows per tour plan with an average lateness of 4.0 minutes. DYN-BUF is able to reduce these figures by accepting two requests less (45.6 vs. 47.2) and enforcing more second time window options than DYN (39.1 vs. 44.0 first options). DYN-SBF reduces the occurrence of lateness through a smaller number of accepted requests (42.1) and significantly less first time window options (33.9). Again, DYN-SBF starts to reject requests relatively sooner, while DYN still accepts more of the later requests. The average distance of rejected customers to the depot is the least for static acceptance mechanisms. Introduction of buffers (DYN-SBF/DYN-BUF) leads to customers being rejected that are farther away from the depot.

The impact of varying customer locations on efficiency and service quality is immense. DYN provides the maximum number of accepted requests and enables the largest number of first time window options. When customer locations are in the suburbs, the number of accepted requests drops by up to 41%, while tour plans become prone to lateness. DYN-BUF helps adjust the occurrence of lateness to a level which might be sufficient from both a service quality and a profit maximization point of view. The intended service level can be adjusted by variation of the parameter α (cf. Section 6.4).

6.2. Varying service times

For the following experiment, instances with varying service times are investigated, inducing varying levels of travel time proportion in tour plans. One would expect that smaller service times would lead to the acceptance of more requests, but also to increased lateness due to an increased proportion of travel time and travel time variation.

In Table 3, results for instances with a service time of 10, 20 and 30 minutes are reported. The number of requests that can be accommodated for 10 minutes service time instances is more than twice as much as for 30 minutes service times. However, the ratios of accepted requests with respect to DYN are within a comparable level (except for DYN-SBF).

For downtown customers, DYN accepts 116.4 requests on average for the 10 minutes instances. Tour plans are so dense that lateness starts to matter even in downtown areas. For DYN, 92% of tour plans run late, but the amount of lateness is still relatively small. Due to the 5-minute buffer, DYN-SBF rejects so many requests that only 75% of the number of DYN is accepted, leading to less denseness and less occurrence of lateness. When service times are extended to 20 and 30 minutes, the occurrence and extent of lateness become negligible for downtown instances, while denseness remains about the same. This is due to the decreasing temporal proportion of travel times in tour plans.

For suburban customers, tour plan distances are again about three times longer than for downtown instances. DYN-SBF again accepts significantly less requests than DYN and about the same as the static mechanisms while reducing lateness heavily. This comes at the cost of first time window options. Lateness occurs noticeably for the 10 minutes instances, where 99% of all DYN

Table 3
Results for varying service times.

Customer location	Acceptance mechanism	# Of pot. feasible requests	# Of posed requests	Service time	Length of time window	Avg # accepted requests	Avg # requests rel. to DYN (%)	Avg # of first options	Avg # of first options rel. to DYN (%)	Tour plan distance	Tour plan denseness (%)	Avg arrival of rej. requests	Avg dist. of rej. customers	Proportion tour plans late (%)	Avg # of tw violations	Avg late at customer	Avg 95th late at customer
Downtown	SLOT	104	130	10	60	103.8	89	93.1	80	160	89	0.85	4.1	36	0.5	0.0	1.4
	TD-SLOT	106				105.7	91	94.8	81	160	90	0.86	4.1	51	0.8	0.0	1.9
	DYN-SBF	–				86.7	75	66.8	57	142	77	0.67	4.3	1	0.0	0.0	1.0
	DYN-BUF(2.0)	–				115.5	99	103.3	89	156	95	0.83	4.3	84	1.8	1.1	3.0
	DYN	–				116.4	100	108.0	93	157	96	0.89	4.2	92	2.4	1.1	3.1
Suburban	SLOT	64	80	10	60	63.6	83	55.7	72	379	79	0.84	8.8	80	1.7	4.1	12.5
	TD-SLOT	66				65.3	85	57.1	74	388	79	0.84	8.7	78	1.5	3.3	9.9
	DYN-SBF	–				62.0	81	48.4	63	371	78	0.67	10.4	65	1.7	6.0	17.8
	DYN-BUF(2.0)	–				72.6	94	61.6	80	407	87	0.70	10.7	93	3.3	5.1	15.5
	DYN	–				76.9	100	72.9	95	424	90	0.91	9.7	99	5.0	5.5	16.4
Downtown	SLOT	56	70	20	60	55.6	90	48.3	78	110	88	0.83	4.1	9	0.1	0.0	0.8
	TD-SLOT	58				57.5	92	49.7	80	113	90	0.83	4.1	22	0.3	0.0	1.4
	DYN-SBF	–				56.5	91	45.5	73	108	89	0.74	4.3	0	0.0	0.0	0.4
	DYN-BUF(2.0)	–				62.0	100	55.4	89	114	95	0.85	4.2	39	0.5	0.0	2.2
	DYN	–				62.2	100	56.6	91	114	96	0.88	4.2	56	0.8	0.0	2.5
Suburban	SLOT	40	50	20	60	39.4	83	33.7	71	284	81	0.81	8.8	59	0.9	4.1	12.0
	TD-SLOT	42				40.8	86	34.7	73	289	82	0.81	8.9	55	0.8	3.3	9.5
	DYN-SBF	–				42.1	89	33.9	72	291	84	0.71	10.2	35	0.6	4.9	14.7
	DYN-BUF(2.0)	–				45.6	97	39.1	83	304	89	0.75	10.3	70	1.4	4.1	12.4
	DYN	–				47.2	100	44.0	93	310	91	0.90	9.5	86	2.3	4.0	12.3
Downtown	SLOT	40	50	30	60	39.3	91	33.3	77	86	89	0.81	4.2	2	0.0	0.0	0.2
	TD-SLOT	42				40.9	95	34.1	79	92	92	0.81	4.2	14	0.2	0.0	1.3
	DYN-SBF	–				41.5	96	35.0	81	92	94	0.80	4.3	0	0.0	0.0	0.3
	DYN-BUF(2.0)	–				43.0	99	37.8	87	93	96	0.85	4.2	20	0.2	0.0	1.7
	DYN	–				43.3	100	38.1	88	93	96	0.86	4.2	33	0.4	0.0	2.1
Suburban	SLOT	32	40	30	60	31.2	89	25.9	73	249	85	0.80	8.7	44	0.6	4.1	12.0
	TD-SLOT	34				32.6	93	27.0	77	250	86	0.80	8.6	55	0.9	3.2	9.2
	DYN-SBF	–				33.3	94	27.1	77	246	89	0.76	9.9	22	0.3	0.0	13.0
	DYN-BUF(2.0)	–				34.5	98	29.5	84	253	91	0.80	9.6	51	0.8	4.1	12.5
	DYN	–				35.2	100	31.6	90	254	92	0.85	9.2	73	1.4	4.4	13.0

Non-overlapping delivery time windows for a typical weekday, 12 pm–8 pm; fleet size = 3 vehicles.

Table 4
Results for varying lengths of service time windows.

Customer location	Acceptance mechanism	# Of pot. feasible requests	# Of posed requests	Service time	Length of time window	Avg # accepted requests	Avg # requests rel. to DYN	Avg # of first options	Avg # of first options rel. to DYN	Tour plan distance	Tour plan denseness	Avg arrival of rej. requests	Avg dist. of rej. customers	Proportion tour plans late	Avg # of tw violations	Avg late at customer	Avg 95th late at customer
Downtown	SLOT	48	60	20	30	45.8	83	37.2	67	102	77	0.77	4.2	16	0.2	0.0	1.2
	TD-SLOT	50				46.8	85	37.2	67	101	74	0.76	4.2	17	0.2	0.0	1.3
	DYN-SBF	–				49.6	90	39.2	71	109	80	0.71	4.4	0	0.0	0.0	0.4
	DYN-BUF(2.0)	–				54.7	99	48.7	88	116	87	0.84	4.3	32	0.4	0.8	2.0
	DYN	–				55.2	100	49.8	90	117	88	0.87	4.3	50	0.7	0.8	2.1
Downtown	SLOT	56	70	20	60	55.6	90	48.3	78	110	88	0.83	4.1	9	0.1	0.0	0.8
	TD-SLOT	58				57.5	92	49.7	80	113	90	0.83	4.1	22	0.3	0.0	1.4
	DYN-SBF	–				56.5	91	45.5	73	108	89	0.74	4.3	0	0.0	0.0	0.4
	DYN-BUF(2.0)	–				62.0	100	55.4	89	114	95	0.85	4.2	39	0.5	0.0	2.2
	DYN	–				62.2	100	56.6	91	114	96	0.88	4.2	56	0.8	0.0	2.5
Downtown	SLOT	64	80	20	120	64.0	96	58.7	88	102	98	0.87	4.1	16	0.2	0.0	1.3
	TD-SLOT	65				65.0	97	59.3	89	105	98	0.87	4.1	14	0.2	0.0	1.3
	DYN-SBF	–				61.7	92	49.4	74	90	94	0.73	4.3	0	0.0	0.0	0.1
	DYN-BUF(2.0)	–				66.8	100	61.5	92	92	99	0.87	4.2	49	0.6	0.0	2.1
	DYN	–				66.9	100	62.2	93	93	99	0.89	4.2	53	0.7	0.9	2.3
Suburban	SLOT	32	40	20	30	30.0	78	23.7	61	251	80	0.74	9.0	63	1.3	4.6	13.8
	TD-SLOT	34				30.6	79	23.7	61	230	62	0.72	9.0	42	0.6	0.0	6.3
	DYN-SBF	–				34.6	90	27.4	71	282	73	0.71	10.0	27	0.5	4.7	14.4
	DYN-BUF(2.0)	–				37.4	97	32.3	84	287	77	0.75	10.7	53	1.0	3.4	11.0
	DYN	–				38.6	100	36.0	93	293	80	0.86	10.4	82	2.2	3.4	11.0
Suburban	SLOT	40	50	20	60	39.4	83	33.7	71	284	81	0.81	8.8	59	0.9	4.1	12.0
	TD-SLOT	42				40.8	86	34.7	73	289	82	0.81	8.9	55	0.8	3.3	9.5
	DYN-SBF	–				42.1	89	33.9	72	291	84	0.71	10.2	35	0.6	4.9	14.7
	DYN-BUF(2.0)	–				45.6	97	39.1	83	304	89	0.75	10.3	70	1.4	4.1	12.4
	DYN	–				47.2	100	44.0	93	310	91	0.90	9.5	86	2.3	4.0	12.3
Suburban	SLOT	52	65	20	120	52.0	93	47.5	85	296	96	0.87	9.0	71	1.2	0.0	13.4
	TD-SLOT	53				53.0	94	48.1	86	298	96	0.87	9.0	67	1.1	4.6	13.2
	DYN-SBF	–				51.7	92	41.8	74	261	94	0.71	10.3	40	0.5	4.4	12.7
	DYN-BUF(2.0)	–				55.6	99	48.7	87	276	97	0.81	9.8	83	1.7	4.6	13.1
	DYN	–				56.1	100	52.4	93	275	98	0.89	9.4	92	2.3	4.9	13.7

Non-overlapping delivery time windows for a typical weekday, 12 pm–8 pm; fleet size = 3 vehicles.

generated tour plans run late with an average 95th lateness of 16.4 minutes. DYN-BUF is able to reduce the occurrence of lateness in a range of 22–65% by accepting less customers than DYN (6–19%), enforcing up to 37% more second time window options. While denseness remains about the same, increasing service times lead to a reduction of the absolute number of accepted requests as well as the occurrence of lateness. Introduction of buffers (DYN-BUF/SBF) again results in rejection of customers that are farther away from the depot.

Variation of service times highlights that more detailed travel time information can be valuable for both downtown and suburban instances, especially when a larger number of customers are serviced. While for downtown instances accepting a slightly smaller number of requests may provide tour plans with sufficient service quality, this is much more challenging in suburban areas, especially for tour plans with small service times and a large number of customers.

6.3. Varying length of service time windows

The length of service time windows is crucial for the perceived service quality as well as for the likelihood of accommodating a request efficiently. We demonstrate the impact of varying time window lengths based on downtown and suburban instances with 30, 60 and 120 minutes service time windows and a fixed service time of 20 minutes. Results are shown in Table 4.

For all instances, the number of accepted requests increases from shorter to longer service time windows, since longer service time windows provide more flexibility to accommodate requests. Static approaches perform relatively better when it comes to longer service time windows.

For downtown instances, SLOT accepts 83% of the number DYN would accept for 30 minutes time windows and 96% of DYN for 120 minutes time windows, for example. Lateness occurrence and extent of lateness remain negligible. Denseness increases from 88% to 96% to 99% for DYN, reflecting that there is almost no idle time due to the flexibility of larger service time windows.

The same is true for suburban customer locations, where static approaches perform relatively better under longer time window conditions. SLOT accepts 78% of DYN for 30 minutes instances and 93% of DYN for 120 minutes instances. For DYN, denseness increases accordingly from 80% to 91% to 98%. This underlines that decisions on the design of service time windows have a large impact on the profitability of delivery tours, especially when customers are located in the suburbs. However, the proportion of late tour plans evolves similarly with the increasing number of accepted requests (82% to 86% to 92%). DYN-BUF is able to reduce the proportion of late tour plans by up to 29% by accepting a slightly smaller number of requests (1–3%).

In summary, longer service time windows enable the acceptance of a significantly larger number of requests. DYN ensures that increasing flexibility can be fully exploited. However, improvements of efficiency are counteracted by decreasing service quality in very dense tour plans. DYN-BUF is able to improve the service quality by slightly reducing the number of accepted requests. We will demonstrate the impact of different settings of α in the next section.

6.4. Impact of lateness avoidance

The idea of lateness avoidance is to reject requests that might be scheduled close to the end of their service time windows, especially in the later part of a dense tour, and thus avoid the occur-

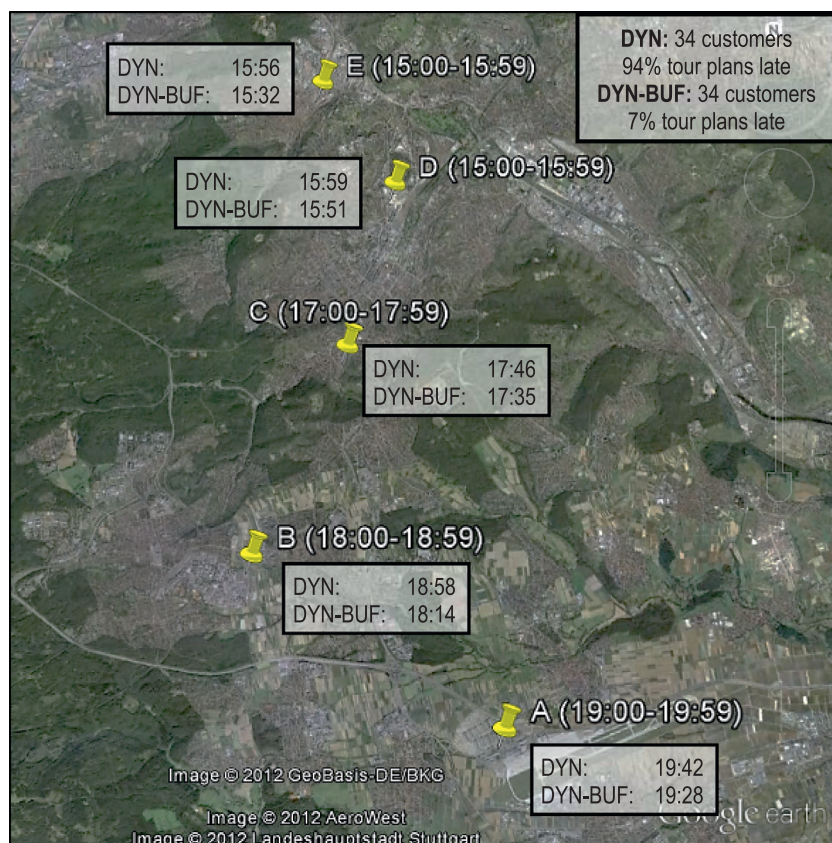


Fig. 4. Example of lateness avoidance. Aerial view provided by Google Earth.

(continued on next page)

Table 5
Results for different settings of α .

Customer location	Acceptance mechanism	# Of posed requests	Service time	Length of time window	Avg # accepted requests	Avg # requests rel. to DYN (%)	Avg # of first options	Avg # first options rel. to DYN (%)	Tour plan distance	Tour plan denseness (%)	Avg arrival of rej. requests	Avg dist. of rej. customers	Proportion tour plans late (%)	Avg # of tw violations	Avg late at customer	Avg 95th late at customer
Downtown	DYN-BUF(2.0)	130	10	60	115.5	99	103.3	89	156	95	0.83	4.3	84	1.8	1.1	3.0
	DYN-BUF(1.0)				116.4	100	105.8	91	156	96	0.86	4.3	85	2.0	1.0	2.9
	DYN-BUF(0.5)				116.2	100	106.7	92	156	96	0.88	4.2	89	2.2	1.1	3.0
	DYN				116.4	100	108.0	93	157	96	0.89	4.2	92	2.4	1.1	3.1
Downtown	DYN-BUF(2.0)	80	20	120	66.8	100	61.5	92	92	99	0.87	4.2	49	0.6	0.0	2.1
	DYN-BUF(1.0)				66.7	100	61.6	92	92	99	0.88	4.2	51	0.7	0.0	2.2
	DYN-BUF(0.5)				66.8	100	62.0	93	93	99	0.88	4.2	52	0.7	0.9	2.3
	DYN				66.9	100	62.2	93	93	99	0.89	4.2	53	0.7	0.9	2.3
Downtown	DYN-BUF(2.0)	70	20	60	62.0	100	55.4	89	114	95	0.85	4.2	39	0.5	0.0	2.2
	DYN-BUF(1.0)				62.0	100	55.9	90	113	95	0.86	4.2	44	0.5	0.0	2.4
	DYN-BUF(0.5)				61.9	100	56.2	90	113	96	0.87	4.2	51	0.7	0.0	2.5
	DYN				62.2	100	56.6	91	114	96	0.88	4.2	56	0.8	0.0	2.5
Downtown	DYN-BUF(2.0)	60	20	30	54.7	99	48.7	88	116	87	0.84	4.3	32	0.4	0.8	2.0
	DYN-BUF(1.0)				54.9	99	49.3	89	116	87	0.85	4.2	41	0.5	0.8	2.1
	DYN-BUF(0.5)				55.0	100	49.5	90	116	87	0.86	4.2	45	0.6	0.8	2.1
	DYN				55.2	100	49.8	90	117	88	0.87	4.3	50	0.7	0.8	2.1
Downtown	DYN-BUF(2.0)	50	30	60	43.0	99	37.8	87	93	96	0.85	4.2	20	0.2	0.0	1.7
	DYN-BUF(1.0)				43.1	100	37.9	88	93	96	0.85	4.2	25	0.3	0.0	2.0
	DYN-BUF(0.5)				43.2	100	38.0	88	93	96	0.85	4.2	28	0.3	0.0	2.0
	DYN				43.3	100	38.1	88	93	96	0.86	4.2	33	0.4	0.0	2.1

Customer location	Acceptance mechanism	# Of posed requests	Service time	Length of time window	Avg # accepted requests	Avg # requests rel. to DYN (%)	Avg # of first options	Avg # of first options rel. to DYN (%)	Tour plan distance	Tour plan denseness (%)	Avg arrival of rej. requests	Avg dist. of rej. customers	Proportion tour plans late (%)	Avg # of tw violations	Avg late at customer	Avg 95th late at customer
Suburban	DYN-BUF(2.0)	80	10	60	72.6	94	61.6	80	407	87	0.70	10.7	93	3.3	5.1	15.5
	DYN-BUF(1.0)				75.1	98	66.6	87	412	88	0.77	10.6	97	4.4	5.7	17.3
	DYN-BUF(0.5)				76.3	99	69.2	90	421	90	0.81	10.2	99	4.7	5.7	17.0
	DYN				76.9	100	72.9	95	424	90	0.91	9.7	99	5.0	5.5	16.4
Suburban	DYN-BUF(2.0)	65	20	120	55.6	99	48.7	87	276	97	0.81	9.8	83	1.7	4.6	13.1
	DYN-BUF(1.0)				55.9	100	50.0	89	276	98	0.84	9.7	89	2.0	4.9	13.7
	DYN-BUF(0.5)				55.9	100	51.2	91	277	98	0.87	9.6	88	2.0	4.9	13.6
	DYN				56.1	100	52.4	93	275	98	0.89	9.4	92	2.3	4.9	13.7
Suburban	DYN-BUF(2.0)	50	20	60	45.6	97	39.1	83	304	89	0.75	10.3	70	1.4	4.1	12.4
	DYN-BUF(1.0)				46.6	99	41.4	88	306	89	0.80	10.1	79	1.8	4.2	12.7
	DYN-BUF(0.5)				47.0	99	42.6	90	308	90	0.84	9.9	83	2.0	4.2	12.8
	DYN				47.2	100	44.0	93	310	91	0.90	9.5	86	2.3	4.0	12.3
Suburban	DYN-BUF(2.0)	40	20	30	37.4	97	32.3	84	287	77	0.75	10.7	53	1.0	3.4	11.0
	DYN-BUF(1.0)				38.1	99	33.9	88	293	79	0.78	10.5	69	1.5	3.4	10.9
	DYN-BUF(0.5)				38.5	100	34.9	91	294	80	0.81	10.4	78	1.9	3.4	10.8
	DYN				38.6	100	36.0	93	293	80	0.86	10.4	82	2.2	3.4	11.0
Suburban	DYN-BUF(2.0)	40	30	60	34.5	98	29.5	84	253	91	0.80	9.6	51	0.8	4.1	12.5
	DYN-BUF(1.0)				34.9	99	30.4	86	255	92	0.82	9.4	66	1.1	4.2	12.7
	DYN-BUF(0.5)				35.1	100	31.0	88	255	92	0.83	9.3	71	1.3	4.3	12.6
	DYN				35.2	100	31.6	90	254	92	0.85	9.2	73	1.4	4.4	13.0

Non-overlapping delivery time windows for a typical weekday, 12 pm–8 pm; fleet size = 3 vehicles.

Table 6

Results for varying customer behavior.

Customer location	Acceptance mechanism	# Of pot. feasible requests	# Of posed requests	Service time	Length of time window	Avg # accepted requests	Avg # requests rel. to DYN (%)	Avg # of first options	Avg # of first options rel. to DYN (%)	Tour plan distance	Tour plan denseness (%)	Avg arrival of rej. requests	Avg dist. of rej. customers	Proportion tour plans late (%)	Avg # of tw violations	Avg late at customer	Avg 95th late at customer
Downtown	SLOT	56	70	20	60	55.6	90	48.3	78	110	88	0.83	4.1	9	0.1	0.0	0.8
	TD-SLOT	58				57.5	92	49.7	80	113	90	0.83	4.1	22	0.3	0.0	1.4
	DYN-SBF	–	–			56.5	91	45.5	73	108	89	0.74	4.3	0	0.0	0.0	0.4
	DYN-BUF(2.0)	–	–			62.0	100	55.4	89	114	95	0.85	4.2	39	0.5	0.0	2.2
	DYN	–				62.2	100	56.6	91	114	96	0.88	4.2	56	0.8	0.0	2.5
Downtown	SLOT	56	70	20	60	50.9	94	40.0	74	100	81	0.74	4.2	3	0.0	0.0	0.4
	TD-SLOT	58				50.8	94	40.0	74	99	80	0.73	4.2	2	0.0	0.0	0.3
	DYN-SBF	–	–			49.1	90	37.6	69	99	84	0.71	4.3	0	0.0	0.0	0.3
	DYN-BUF(2.0)	–	–			53.5	98	44.0	81	101	87	0.77	4.2	40	0.5	0.0	2.1
	DYN	–				54.4	100	45.0	83	103	88	0.77	4.2	53	0.7	0.0	2.3
Suburban	SLOT	40	50	20	60	39.4	83	33.7	71	284	81	0.81	8.8	59	0.9	4.1	12.0
	TD-SLOT	42				40.8	86	34.7	73	289	82	0.81	8.9	55	0.8	3.3	9.5
	DYN-SBF	–	–			42.1	89	33.9	72	291	84	0.71	10.2	35	0.6	4.9	14.7
	DYN-BUF(2.0)	–	–			45.6	97	39.1	83	304	89	0.75	10.3	70	1.4	4.1	12.4
	DYN	–				47.2	100	44.0	93	310	91	0.90	9.5	86	2.3	4.0	12.3
Suburban	SLOT	40	50	20	60	35.6	88	27.4	67	249	73	0.74	8.7	46	0.7	0.0	11.1
	TD-SLOT	42				35.5	87	27.5	68	245	70	0.73	8.6	33	0.4	0.0	7.7
	DYN-SBF	–	–			37.1	91	28.4	70	260	83	0.71	9.6	36	0.6	4.8	14.2
	DYN-BUF(2.0)	–	–			39.5	97	31.8	78	266	84	0.75	9.5	67	1.2	4.4	13.7
	DYN	–				40.7	100	34.6	85	271	85	0.80	9.2	86	2.0	4.8	14.1

Higher demand in the afternoon peak hours.

rence of lateness. In the following, we discuss the effectiveness of DYN-BUF with emphasis on the impact of the parameter α on lateness avoidance.

A sample of five customers from a suburban instance is shown in Fig. 4, including service time windows and mean arrival times provided by DYN and DYN-BUF. For this instance, DYN accepts 34 requests in total, leading to a proportion of 94% late tour plans. This emerges from the fact that scheduled arrival times are often close to the end of service time windows, facing lateness in tour execution. For instance, the start of service for customer B is scheduled to 18:58 by DYN, which is very close to its latest possible start of service at 18:59. The solution provided by DYN-BUF with $\alpha = 2.0$ also yields the acceptance of 34 customers in total, but the start of service for customer B is scheduled to 18:14, which is well before its latest start of service. Thus, simulation of DYN-BUF generated tour plans results in a significantly smaller proportion of late tour plans. Unfavorable requests have either been replaced by second time window options or they have been rejected and replaced by later requests. For the shown customers, it is obvious that the DYN-BUF generated arrival times are much more reliable than the DYN generated arrival times, while the level of efficiency remains constant.

We now vary the parameter α in order to demonstrate its effect on the number of accepted requests and resulting service quality. Table 5 shows the results for all aforementioned instances with $\alpha = \{0.5, 1.0, 2.0\}$. For the downtown instances, the number of accepted requests increases slightly with decreasing alpha. It is always very close to the number of requests accepted by DYN. The proportion of tour plans late decreases with an increasing value of α by an extent of 4–18%. The occurrence of lateness is heavily reduced, extent of lateness decreases slightly. Reduction of lateness occurrence is mainly achieved by enforcement of second time window options (up to 4% more than for DYN), especially when it comes to tour plans with a large number of customers. The higher the value of α , the sooner requests are rejected. Distances of rejected customers to the depot remain about the same.

For the suburban instances, the same variation of α leads to larger differences in the number of accepted requests as well as in the proportion of late tour plans. With α varying between 0.5 and 2.0, 94–100% of the maximum number of requests is accepted, enforcing up to 15% more second time window options than DYN. The proportion of late tour plans is reduced by 6–29%, accompanied by a reduction of the average number of violated service time windows. While with decreasing value of α the average arrival of rejection becomes later, the average distance of rejected customers to the depot decreases (up to 1 km). Thus, with increasing buffers, customers seem to be avoided that require a larger detour.

A comparison of the suburban instances underlines the adaptability of DYN-BUF to customer locations, length of service time windows and denseness of tour plans. In general, DYN-BUF with $\alpha = 2.0$ is able to accept as nearly as many requests as DYN and significantly more than static approaches. The average proportion of late tour plans is smaller, for some instances even as small as the proportion of static approaches. DYN-SBF yields a lower number of accepted requests and first time window options, but also reduces the occurrence of lateness significantly.

Overall, when customers are located in suburban areas, static approaches accept significantly less than DYN while still leading to some occurrence of lateness. DYN provides the maximal number of accepted requests, accompanied by a maximum level of denseness and occurrence of lateness. DYN-BUF is able to balance the number of requests and resulting service quality. The proportion of late tour plans is always lower than in the DYN case, and it is often even close to the occurrence of lateness induced by more conservative static approaches. For downtown customer locations, the same is true in general. However, the extent of lateness is often

negligible since tour plans are characterized by small distances. Here, DYN-SBF may often eliminate lateness, but it comes with a heavily reduced number of first time window options, which is sometimes even lower than the number produced by static acceptance mechanisms.

6.5. Varying customer behavior

For the previous experiments, it was assumed that customer demand for all the offered time slots is uniform. We now demonstrate the impact of more realistic customer behavior on the effectiveness of the proposed acceptance mechanisms. For this experiment, we suppose that customers might prefer delivery time slots in the late afternoon, when most people get home from work. We have adjusted the probabilities such that time slots between 4 pm and 6 pm are much more likely selected than the remaining ones. Results are presented in Table 6 with shadowed areas denoting the instances referring to the modified customer behavior.

Higher demand in afternoon hours leads to a drop in the number of accepted requests (DYN: about 14%) and an even larger decrease of first time window options (DYN: about 26%), accompanied by smaller levels of denseness. The superiority of TD-SLOT over SLOT diminishes, since demand for service in rush hours increases, where TD-SLOT enforces a smaller maximum number of service requests.

Higher demand in afternoon hours demonstrates the typical challenges of attended home delivery in metropolitan areas. When demand for service is at its peak, traffic conditions tend to be worse. Only dynamic approaches are able to determine more cost-efficient, but still feasible tour plans. However, profitability drops significantly compared to uniformly distributed demand for time slots because of the preference for service time windows with increased travel times. These tests underline the necessity of customer incentives in the management of delivery time windows.

7. Conclusions and outlook

We have discussed challenges arising from the acceptance of delivery requests in congested metropolitan areas. Static and dynamic acceptance mechanisms were evaluated by computational experiments. Computational experiments have provided insights on how additional travel time information can improve profitability and feasibility of tour plans. They can be summarized as follows:

- It is important to carefully consider the location of customers in the management of service time windows. While travel time variation does not play a significant role for close downtown customer locations, congestion on arterial roads has a significant impact on the profitability and the service quality of attended deliveries.
- Consideration of travel time evolution improves the number of accepted requests for static acceptance mechanisms. Time-dependent acceptance mechanisms outperform their non-time-dependent counterparts by estimating logistics capabilities more precisely.
- A dynamic acceptance mechanism provides the largest profits for all instances. This is because it is able to adapt to a specific set of customer requests, including time-dependent travel time information arising from a potential tour plan.
- For close customer locations in downtown areas, a basic dynamic acceptance mechanism provides solutions that are of high profitability and service quality, while for customer locations in suburban areas, tour plans are prone to lateness.

- A simple buffer approach may be suitable to reduce lateness for downtown areas. However, a fixed buffer cannot sufficiently reflect dynamic traffic conditions and varying customer locations in suburban areas. A fixed buffer also comes with the cost of rejecting more of the first priority customer time window options.
- Additional information on travel time variation is very valuable in order to alleviate lateness and lateness propagation. This is especially important for short service time windows and suburban customer locations.

So far, our approach can be classified as a post-processing feasibility check of a time-dependent routing solution. Future research suggests the consideration of lateness propagation in the actual construction of a TDVRPTW solution in order to further increase the number of accepted requests while ensuring a certain level of service quality. The challenge here is to avoid costly recomputation of buffers for all investigated tour variants. Furthermore, analysis of more realistic customer behavior based on real order data and the incorporation of customer incentives could provide valuable insights and improve profitability of attended home deliveries.

Acknowledgement

This work was supported by a fellowship within the Postdoc-Programme of the German Academic Exchange Service (DAAD, Jan F. Ehmke).

References

- Agatz, N., Campbell, A. M., Fleischmann, M., & Savelsbergh, M. (2011). Time slot management in attended home delivery. *Transportation Science*, 45(3), 435–449.
- Ando, N., & Taniguchi, E. (2006). Travel time reliability in vehicle routing. *Networks and Spatial Economics*, 6(3), 293–311.
- Azi, N., Gendreau, M., & Potvin, J.-Y. (2012). A dynamic vehicle routing problem with multiple delivery routes. *Annals of Operations Research*, 199(1), 103–112.
- Baldacci, R., Mingozzi, A., & Roberti, R. (2012). Recent exact algorithms for solving the vehicle routing problem under capacity and time window constraints. *European Journal of Operational Research*, 218(1), 1–6.
- Bräysy, O., & Gendreau, M. (2005a). Vehicle routing problem with time windows, Part I: Route construction and local search algorithms. *Transportation Science*, 39(1), 104–118.
- Bräysy, O., & Gendreau, M. (2005b). Vehicle routing problem with time windows, Part II: Metaheuristics. *Transportation Science*, 39(1), 119–139.
- Campbell, A. M., & Savelsbergh, M. (2005). Decision support for consumer direct grocery initiatives. *Transportation Science*, 39(3), 313–327.
- Campbell, A. M., & Savelsbergh, M. (2006). Incentive schemes for attended home delivery services. *Transportation Science*, 40(3), 327–341.
- Chang, T.-S., Nozick, L. K., & Turnquist, M. A. (2005). Multiobjective path finding in stochastic dynamic networks with application to routing hazardous materials shipments. *Transportation Science*, 39(3), 383–399.
- Chang, T.-S., Wan, Y.-W., & Ooi, W. T. (2009). A stochastic dynamic traveling salesman problem with hard time windows. *European Journal of Operational Research*, 198(3), 748–759.
- Donati, A. V., Montemanni, R., Casagrande, N., Rizzoli, A. E., & Gambardella, L. M. (2008). Time dependent vehicle routing problem with a multi ant colony system. *European Journal of Operational Research*, 185(3), 1174–1191.
- Ehmke, J. F., Meisel, S., & Mattfeld, D. C. (2012b). Floating car based travel times for city logistics. *Transportation Research Part C: Emerging Technologies*, 21(1), 338–352.
- Ehmke, J. F., Meisel, S., & Mattfeld, D. C. (2010). Floating car data based analysis of urban travel times for the provision of traffic quality. In J. Barceló & M. Kuwahara (Eds.), *Traffic data collection and its standardization* (pp. 129–150). New York: Springer.
- Ehmke, J. F., Steinert, A., & Mattfeld, D. C. (2012a). Advanced routing for city logistics service providers based on time-dependent travel times. *International Journal of Computational Science*, 3(4), 193–205.
- Fleischmann, B., Gietz, M., & Gnuzmann, S. (2004). Time-varying travel times in vehicle routing. *Transportation Science*, 38(2), 160–173.
- Haghani, A., & Jung, S. (2005). A dynamic vehicle routing problem with time-dependent travel times. *Computers & Operations Research*, 32(11), 2959–2986.
- Hashimoto, H., Yagiura, M., & Ibaraki, T. (2008). An iterated local search algorithm for the time dependent vehicle routing problem with time windows. *Discrete Optimization*, 5(2), 434–456.
- Ichoua, S., Gendreau, M., & Potvin, J. Y. (2003). Vehicle dispatching with time-dependent travel times. *European Journal of Operational Research*, 144(2), 379–396.
- Kim, S., Lewis, M. E., & White, C. C. III. (2005). Optimal vehicle routing with real-time traffic information. *IEEE Transactions on Intelligent Transportation Systems*, 6(2), 178–188.
- Kok, A. L., Hans, E. W., & Schutten, J. M. J. (2012). Vehicle Routing under time-dependent travel times: The impact of congestion avoidance. *Computers & Operations Research*, 39(5), 910–918.
- Lau, H. C., Sim, M., & Teo, K. M. (2003). Vehicle routing problems with time windows and a limited number of vehicles. *European Journal of Operational Research*, 148(3), 559–569.
- Maden, W., Eglese, R., & Black, D. (2010). Vehicle routing and scheduling with time-varying data: A case study. *Journal of Operations Research Society*, 61, 515–522.
- Pillac, V., Gendreau, M., Guéret, C., & Medaglia, A. L. (2013). A review of dynamic vehicle routing problems. *European Journal of Operational Research*, 225(1), 1–11.
- Pisinger, D., & Ropke, S. (2007). A general heuristic for vehicle routing problems. *Computers & Operations Research*, 34(8), 2403–2435.
- Punakivi, M., & Saranen, J. (2001). Identifying the success factors in e-grocery home delivery. *Journal of Retail and Distribution Management*, 29(4), 156–163.
- Schonfeld, E. (2010). Forrester forecast: Online retail sales will grow to \$250 billion by 2014. <http://www.techcrunch.com/2010/03/08/forrester-forecast-online-retail-sales-will-grow-to-250-billion-by-2014/>.
- Solomon, M. M. (1987). Algorithms for the vehicle routing and scheduling problem with time window constraints. *Operations Research*, 35, 54–65.
- Susilawati, S., Taylor, M. A. P., & Somenahalli, S. V. C. (2011). Distributions of travel time variability on urban roads. *Journal of Advanced Transportation*. <http://dx.doi.org/10.1002/atr.192>.
- TomTom (2012). Real time & historical traffic. <<http://www.tomtom.com/lib/doc/licensing/RTTHT.EN.pdf>> Accessed 01.02.13.
- US Census Bureau (2012). Quarterly Retail E-Commerce Sales 2nd Quarter 2012, news release. <http://www.census.gov/retail/mrts/www/data/pdf/ec_current.pdf> Accessed 16.08.12.