OR Spectrum (2010) 32:831–859
DOI 10.1007/s00291-009-0190-7

REGULAR ARTICLE

Adaptive demand peak management in online transport process planning

Jörn Schönberger

Published online: 22 December 2009

Abstract We investigate the common and integrated dynamic decision making of the coordinator of a supply chain consortium together with a subordinate fleet managing agent offering transport services for the consortium. While the fleet manager aims at minimizing the costs of the generated transport processes, the goal of the coordinator is to keep the reliability and stability of the processes on a reasonable level. It aims to synchronize the transport processes with upstream and downstream parts of the supply chain. The major innovation presented in this article is a framework that controls and adjusts the decision competence distribution between the two planning agents with respect to the current transport process performance. If the transport process timeliness is endangered to fall below a given threshold and thereby the overall supply chain reliability tends to sink, the coordinator is temporarily granted the right to intervene into the planning of the fleet managing agent. Within simulation experiments, we demonstrate that the proposed system is able to increase the reliability of the generated transport processes. We show that the intervention of the superior coordinator agent during workload peaks ensures higher process timeliness than the transport service providing agent is able to achieve without any coordinator interventions.

Keywords Dynamic decision problem · Transportation · Online optimization · Adaptation · Principal–agent relationship

1 Introduction

A supply chain is a temporal coalition of specialized partners that is formed in order to execute collaboratively all necessary value creation steps of a specific product.

J. Schönberger (✉)
Chair of Logistics, University of Bremen, Wilhelm-Herbst-Straße 5, 28359 Bremen, Germany
e-mail: jsb@uni-bremen.de

All partners contribute to the success of the coalition. One partner in the coalition (the coordinator) has been granted the right to coordinate the activities of the other partners (service agents). This competency distribution induces an intuitive hierarchy among the coordinator and each involved service-providing agent. This hierarchy is named principal–agent relationship (Jensen and Meckling 1976). The process planning in principal–agent relationships in general and in supply chain in special is compromised by organizational borders among the partners (Kaluza et al. 2003). Each partner wants to preserve and protect its own decision competence in the supply chain consortium and wants to keep its autonomy. However, an integration of the decision making of the coalition members is necessary to coordinate the value-creation processes through the different value creation-stages managed by the coalition members.

Transport enables the integration of spatially distributed procurement, production, and distribution as well as sales facilities. It is therefore a core service for supply chain coalitions. The planning of efficient and effective transport processes is compromised by consecutively released information about additional demand, resource availability, or process quality. A recurrent revision of once fixed transport processes is necessary. In this context, a major challenge in operational transport process planning is the management of spontaneously and temporarily occurring load peaks. Peak management concepts applied in other value creating fields, e.g., production, cannot be transferred due to their non-consideration of the spatial dimension of transport planning.

In this article, we report about investigations to increase the transport process reliability (timeliness) during and immediately after demand peaks. The basic idea followed in this article is to enable the supply chain coordinator to adaptively intervene in the resource allocation carried out by the transport service-providing agent. We aim at proving (or disproving) the following research hypothesis: *Granting a supply chain coordinator temporarily the right to allocate transport resources of a subordinate service agent supports increasing the timeliness of the generated supply chain wide transport processes*.

The structure of this article is as follows: Sect. 2 is dedicated to the description of the transport process planning in supply chain scenarios. A special decision scenario from transportation is introduced in Sect. 3. Section 4 reports the configuration of a rolling horizon planning system with adaptive coordinator intervention and Sect. 5 contains the results observed in numerical simulation experiments.

2 Dynamic transport process planning in supply chain coalitions

We start with the provision of a vocabulary to describe re-planning problems (Subsect. 2.1). Next, we discuss the derivation of transport processes in a supply chain consortium (Subsect. 2.2) and explain how these transport processes are endangered by uncertain planning data (Subsect. 2.3). The resulting request management requirements are exposed (Subsect. 2.4). Finally, we discuss ideas to manage workload peaks (Subsect. 2.5).

2.1 Classification of decision problems and decision models

A decision problem P is called *static* if it can be solved completely at one time point t_0 (Sellmaier 2007). In all other cases, a decision problem is called a *dynamic problem* or an *online* (*decision*) *problem*. Such a decision problem is characterized by the need of a repeated decision making and decision revision at consecutive times $t_0, t_1, t_2, \ldots$ (Brehmer 1992). Consequently, a dynamic decision problem P is equivalent to a sequence $P_0, P_1, P_2, \ldots$ of concatenated static decision problems. The decision problem P_{i+1} represents P_i enriched by the additional data acquired between t_i and t_{i+1}. An online decision problem is referred to as *real-time decision problem* if the time to re-compute a solution (deliberation time) of the instance P_i of the online problem is limited (Séguin et al. 1997). Decision problems are often dynamic because decision-relevant information is missing at a decision time t_i (Lund et al. 1996). Typically, decisions made at time t_i have impacts to the decision alternatives available in later decision situations $t > t_i$ (Busemeyer 2001). If the appearance of future events can be approximated by probability definitions, then a dynamic decision problem is called a *Markov Decision Process* (Littman and Dean 1995). The portion of information revealing after the initial decision time t_0 is expressed by the *degree of dynamism* (Lund et al. 1996; Larsen et al. 2008). In transportation logistics, dynamic assignment problems (McKendall and Jaramillo 2006; Powell 1996; Powell and Cheung 2000; Fleischmann et al. 2004; Mitrović-Minić and Krishnamurti 2004; Pankratz 2002; Lund et al. 1996; Madsen et al. 1995) as well as combined dynamic vehicle routing and scheduling problems [dynamic vehicle dispatching problems (Rego and Roucairol 1995)] are major research fields. Furthermore, dynamic emergency team relocation problems (Rajagopalan et al. 2008) are in the focus of scientific interest. A survey of the differences between static and dynamic vehicle deployment is given in Psaraftis (1995, 1988). If all data that are considered for the definition of the decision problem P are assumed to be definite, then P is called a *deterministic* (Brage Illa 1966) decision problem; otherwise, P is named *non-deterministic* (Munera 1984).

Let M be a decision model for the decision problem P. The model M is called a *static model* if the realization time of a decision is irrelevant for the feasibility and the evaluation of the decision (Sellmaier 2007; Saéz et al. 2008). The mixed integer linear program formulation of the traveling salesman problem (Garfinkel 1985) is a representative example for a static model: Only the originating and the terminating nodes determine the costs for traveling along an arc but not the time when the arc is traversed. If a model M is not a static model, then M is called a *dynamic model* (Sellmaier 2007) that is applied if, e.g., capacity availabilities or realization costs alter over time. Dynamic optimization models are used to represent routing tasks if time or costs to pass through roads or paths alter over time (Fleischmann et al. 2004; Moreira et al. 2007; Potvin et al. 2006; Ichoua et al. 2003). Other applications of dynamic models comprise (among others) core control in computer sciences (Ramamritham and Stankovic 1994) or machine scheduling (Seiden 1996). If uncertain data are approximated by probability distributions, then M is called a *stochastic decision model* (Laux 1982). Let M_i denote a decision model for the instance P_i of an online decision model. The sequence of decision models $M_0, M_1, \ldots$ is called an *online decision model* for the online decision problem $P_0, P_1, \ldots$.

2.2 Transport process planning in a supply chain consortium

A supply chain (SC) is a temporal coalition of several value creating partners which are independent by organization and/or by law. Each partner of the coalition provides highly specialized and skilled competencies in a particular stage of the creation of a particular good or class of goods. The coalition partners provide their services at spatially scattered locations so that transport is a necessary and important service in order to bridge the distances between the different facilities forming the physical supply chain.

The coordinating unit of the SC (*coordinator agent*) is responsible for the achievement and maintenance of a high-order fulfillment quality and reliability. A second decision making unit is the fleet manager (*fleet management agent*). It is a profit-center-like unit in the SC-consortium that receives a budget to fulfill the necessary transport requests. The profit of the fleet manager is the difference of the budget (available for a certain period) and the spent costs. In order to achieve a maximal profit, the fleet management agent aims at generating highest profitable transport processes from the transport requests taking into account planning guidelines and parameters fixed by the coordinator agent. Its planning task requires the solving of a combined routing and scheduling problem with subcontracting (Krajewska and Kopfer 2009; Schönberger 2005; Pankratz 2002; Feillet et al. 2005). Typically, other profit centers participate in the SC coalition, but we restrict ourselves to the analysis of the cooperation of the two aforementioned decision making agents.

Both planning agents interact in order to determine the necessary transport processes for the fulfillment of the customer order fulfillment tasks. Especially, they interchange different kinds of information with each other. The recent quality of the generated processes (e.g., the current punctuality quote p_t at a time t) is fed back to the coordinator agent (step A in Fig. 1). This agent has to decide if the so far used planning guidelines and parameters can be kept or if they require an update. In case that an adjustment of the planning premises is identified as necessary or beneficial by the coordinator agent, it submits the updated guidelines to the fleet management agent (step B in Fig. 1).

Different planning objectives (goals) are addressed by the two decision making units. The coordinator agent has the right to intervene into the fleet manager's planning behavior (by the specification of the planning guidelines and the determination of planning parameters), but the fleet management agent decides about the final dispatching of the transport resources in the limits set by of the current planning guidelines and parameters. Therefore, the task of determining transportation process planning task introduced before is predestined to be solved by a multi-agent system (Woolridge 2002).

The derivation of profitable and reliable transport processes is as follows. Customers contact the coordinator agent in order to express and submit their demand to get fulfilled. The customer's demand is expressed as an (external or customer) order (step 1 in Fig. 1) towards the supply chain coordinator. Several external orders (often of different customers) are consolidated by the coordinator which then derives internal requests to be executed at the different SC-stages. The coordinator then allocates resources for fulfilling the internal requests at the supply chain partners (step 2). Especially, transport requests are specified and handed over to the fleet managing agent (step 3). The fleet manager consolidates the transport requests into executable

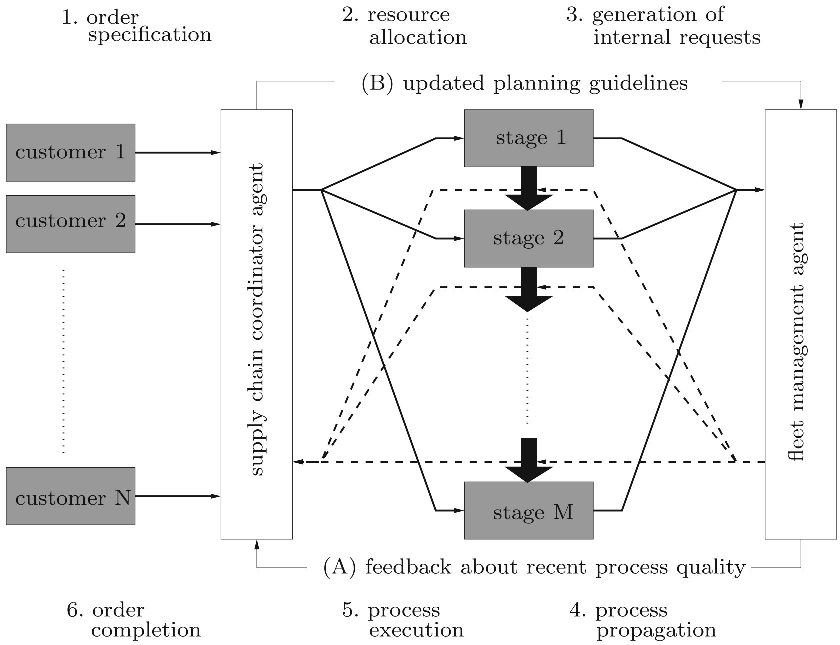

Fig. 1 Supply-chain-wide transport dispatching by a two-agent-system

transport processes which are propagated to the transport providing resources (step 4). The execution of the transport processes contributes to the fulfillment of the internal requests (step 5) whose execution then leads to the completion of the external customer orders (step 6).

A *transportation plan* describes how the known requests are fulfilled (Crainic and Laporte 1997). Subsequently arriving requests are accepted and handled by updating the existing transportation plan. Therefore, the planning situation outlined above is a dynamic decision situation. Each instance P_i of this problem is static and deterministic (all data relevant at the re-planning time t_i are assumed to be known).

2.3 Transport process endangering events and data uncertainty

The compilation of (freight) transport processes is a very challenging dynamic decision task which requires the solving of integrated selection, assignment, and selection problems (Krajewska and Kopfer 2009; Schönberger 2005). Three types of unexpected or unpredictable events endanger and compromise the complete execution of the found processes (Zeimpekis et al. 2007).

Portfolio modifications. Changes in the maintained portfolio of requests are the major driver for a transport process update. Additional requests must be integrated into previously derived transport processes (Mitrović-Minić and Krishnamurti 2004; Ausiello et al. 1994; Lund et al. 1996; Saéz et al. 2008), cancellation of requests sometimes cause a process adaptation in order to keep or re-achieve the process profitability (Rego and Roucairol 1995), and incomplete or wrong predictions of the actually demanded transport resource capacity let an update of the once fixed processes become necessary if the once allocated capacity is exceeded by the additional demand (Lund et al. 1996).

Transport resource shortages. Vehicle breakdowns and vehicle outages shrink the capacity of available transport resources so that some requests cannot be picked up and moved as planned (Zeimpekis and Giaglis 2005).

Infrastructure modifications. Events restricting the infrastructure availability form the third group of incidences compromising the execution of once planned transport processes. Traffic congestion and adverse traffic conditions prolonging the travel times along roads are in the focus of scientific research (Lo and Hall 2008; Fleischmann et al. 2004).

Uncertainty is caused by the occurrence of the aforementioned events and refers to the vagueness of the data used to decide about the deployment necessary to fulfill the current request portfolio. *Request Uncertainty* refers to spontaneous changes of the portfolio of requests (appearance of additional customer requests, request cancellation, request variation) considered so far in the processes (Fleischmann et al. 2004; Madsen et al. 1995). Delays of down-stream activities during the production or preparation of the goods associated with transport requests often cause delays in subsequent value creation stages, e.g., a delayed provision of the goods for the transportation (*Downstream Uncertainty*) (Williams 1984). Last-minute changes of the quantities to be moved are critical if the difference between the ex ante announced and the actual volume to be moved is not available on the selected vehicle (*Quantity Uncertainty*). *Handling Uncertainty* refers to on-site operations of loading or unloading which lasts longer (or shorter) than planned (Hadjiconstantinou and Roberts 2002) and *Transshipment Uncertainty* expresses possible delays caused by the non-availability of ramps, gates, or special loading or unloading equipment in transshipment or cross-docking terminals. If different types of goods cannot be loaded as planned or are not allowed to be loaded by the same resource, then *Loading Uncertainty* (Delaitre et al. 2007; Gendreau et al. 2004) compromises the process execution. The movement of picked up goods from the loading place to the unloading location is compromised by blocked or congested roads so that the vehicle travel time is prolonged (*Travel Uncertainty*). Consequently, a late arrival at the delivery place results (*Upstream Uncertainty*) (Lo and Hall 2008). Often, one kind of event causes another kind of process compromising events so that several sources of uncertainty have to be handled simultaneously. Independently from the type of uncertainty, the incomplete knowledge of the exact planning data leads to the need for revisions of routing and scheduling decisions.

2.4 Request management challenges in dynamic dispatching

The repeated update of transport processes comes along with several special requirements. It is important that once made decisions can be protected, e.g., they cannot be revised. In transportation dispatching not all decisions can be boundlessly revised. The reliability of announcements towards customers concerning arrival times are an important and distinguishing service aspect so that the associated decisions should be left unchanged during process revisions (Madsen et al. 1995; Erera et al. 2008). However, some decisions are even unable to be revised at all like the decision to accept a request.

The urgency of the requests in a portfolio R_t has to be considered. Requests whose completion is necessary in the near future have to be treated preferentially compared to requests whose completion is expected to be far away from the current time (Mitrović-Minić and Krishnamurti 2004). If some urgent requests cannot be handled immediately, then it is necessary to manage the resulting waiting list or queues efficiently (Fleischmann et al. 2004). In case those late customer site arrivals cannot be prevented, delays have to be handled so that the negative impacts of unsatisfied customer requirements are minimized (Fleischmann et al. 2004).

There is a time gap between the decision how a request is executed and the realization (Lund et al. 1996). From the decision time until the realization time the value of a decision might vary (Albers et al. 2001), e.g., the appropriateness of a decision can be changed so that decision revisions become necessary.

Clearly, general challenges in dynamic decision making also apply for dynamic transport resource dispatching. At first, Each decision made at time t_i influences the number of decision alternatives and their evaluation for later re-planning times $t_{i+1}, t_{i+1}, \ldots$ (Saéz et al. 2008). Anticipation of future system states is necessary (Branke and Mattfeld 2005). Furthermore, In dynamic decision making each solution update follows one or more given global planning objectives that guide the decision making over the consecutive update stages. Lund et al. (1996) proposes to deviate temporarily from the followed superior decision strategies by varying objectives and decision alternatives in order to heal quicker negative impacts of extraordinary events. This last topic in the previous list will be particularly addressed in the remainder of this article.

2.5 Operational demand peak management in vehicle dispatching

The generated transport processes specify the transport services to be carried out. Pickup locations, delivery locations, quantities to be moved, and operation times are determined so that the deployed transport resources are used with highest efficiency. However, the completion of the processes is endangered to be compromised by uncertainty resulting from the occurrence of unforeseen events (e.g., additional requests) as discussed earlier.

We investigate the specific situation in which besides the time and location of the additional requests, the number of additional requests released at a certain time is not known in advance. Actually, a balanced stream of regularly incoming requests is replaced by a temporary workload peak that begins at a previously unknown time and lasts an unknown period. During the peak period, the number of incoming requests at a certain time is significantly higher than before and after the peak period.

If the capacity of the available transport resources is already exhausted by the so-far-known requests, then the additional requests cannot be integrated into the processes. The flexibility of the processes decreases (Schönberger and Kopfer 2009a) which means that the probability for serving the additional requests in compliance with the customer time window requirements decreases. In order to ensure that the flexibility of the generated processes remains on a satisfying level even if the workload increases spontaneously, additional capacity must be provided immediately after the peak's occurrence.

Obviously, the reservation (blockage) of *emergency capacity* for the vehicles belonging to the controlled fleet supports the compensation of workload peaks. However, two aspects compromise this remedying idea. (1) The additional resources are not profitable during off-peak periods. Therefore, it is not possible to maintain such a security capacity over a longer period. (2) If none of the vehicles with emergency capacity is located in the nearer surrounding of the request-associated location, then additional capacity reserved only for peak situations is worthless.

How to cope with an overloading of the own fleet? The acceptance and exploitation of shortfalls and under-capacity situations is proposed (Zäpfel 1982) but, due to the currently customer-dominated transportation market, not enforceable. Also temporal staffing arrangements (temporary workers) (Zäpfel 1982; Kalleberg et al. 2003) cannot remedy transportation capacity bottlenecks since additional vehicles must be available, too. Additional deliveries and postponement are also proposed (Zäpfel 1982) but, again, not applicable in freight transportation logistics with supply functions since in-time and complete deliveries are necessary to keep the SC performance alive. A deferred delivery corrupts the flow of goods along the SC. The spatial pre-distribution of deliverable goods over the operations area and a pre-loading of trucks is proposed and evaluated in Calza and Passaro (1997). However, these approaches are very capital-consuming. Pre-peak resource build-ups (Ronen et al. 2001) are also impossible since the peak occurrence is not predictable and transport services cannot be stored. These previously mentioned strategies to mitigate the negative impacts of demand peaks have their origin in production management and exploit the special conditions in this sector. Consequently, the strategies cannot be transferred into the completely different area of transportation.

Instead of holding or blocking a certain amount of capacity constantly in readiness, a temporal and demand-oriented provision of additional transport capacity is preferable (external procurement, third party supply, externalization). External forwarders or transport service providers are incorporated in order to execute selected requests (subcontracting). These logistic service providers (LSPs) provide capacity in their transport resources (e.g., vehicles). In order to ensure that the subcontracting opportunity is available whenever and wherever it is needed, the coordinator sets up and maintains longer-term contracts with one or more service-providing companies covering the area served by the considered SC (Ronen et al. 2001). These contracts specify the provided capacities and the fees which are paid to the service providers. A survey on subcontracting arrangements can be found in Kopfer et al. (2007).

In the remainder of this research we assume that the coordinator agent and a sufficiently large number of logistic service providers have agreed respective contracts on a request-oriented base (Kopfer et al. 2007; Krajewska and Kopfer 2009).

According to the contract, each single request can be given to an LSP. The decision-making units in the SC (the coordinator and/or the fleet managing agent) have to decide which requests are subcontracted. We further assume that exactly one LSP can be chosen to fulfill a certain request so that no provider selection becomes necessary. The service provider starts immediately with the execution of the request and ensures a timely fulfillment. We refrain from explicitly considering preparation or setup times of the LSP. A previously fixed fee is paid by the coordinator agent to the LSP for each selected request.

The contracts with the LSPs secure the SC-consortium an opportunity to make use of the resources of the LSP if necessary. Thereby, the coordinator agent as well as the fleet managing agent have the *right but not the obligation* to realize this choice. For this reason, the externalization (subcontracting) of a request is interpreted as an *option*. The SC-consortium (by means of the coordinator or the fleet managing agent) can exert this option at every time during the duration of the contract (*American option*). Sometimes, the option remains unused. The price for exercising the option is fixed at the time when the contract is signed (*call option*) (Black and Scholes 1973).

Prices to be paid to the LSP in case those options are used are quite high, because the prices must include the costs for hedging the LSP's risk that the SC-consortium does not exercise an option. If both the SC-owned resources and the options are available, then the first-mentioned alternative would be preferentially selected by the SC-consortium. However, if the SC-owned resources are exhausted, then the higher expenses for LSP-incorporation are justified by the timely request execution that supports the goal to keep the quote of reliably completed customer demand high.

3 A dynamic vehicle routing problem in a supply chain setting

We introduce a transport process planning scenario in which the supply chain coordinator and the transport service agent interactively cope with a workload peak caused by an unpredicted temporary increase of additionally appearing customer demand. We scrutinize the induced dynamic decision situation from the viewpoint of the coordinator of the supply chain coalition (Subsect. 3.1) as well as from the position of the subordinate transport service agent (Subsect. 3.2).

3.1 The coordinator agent's decision task

The principal contribution of the coordinator agent to the dispatching of the fleet is to scan the incoming requests and to select requests for which an LSP-option is exercised. In order to keep the once generated processes as stable as possible and in order to reduce the nervousness of the processes (Schönberger and Kopfer 2008) the number of options exercised for already scheduled requests should be kept as small as possible. Only for additionally arrived requests which have not yet been initially scheduled it is beneficial to decide about the exercise of LSP-options. The set $R^{+}_{t_i}$ comprises all requests released at time t_i and the decision problem P_i^{coord} of the SC-coordinator is now to

1. determine an adequate portion f_{t_i} of the requests in $R^+_{t_i}$ for which the LSP-options are exercised (realized) at the re-planning time t_i (*intensity determination*).
2. select the previously fixed number of requests from $R^+(t_i)$ and to put them into $R(t_i, f_{t_i})$ (*request choice*).

The coordinator agent's decision task is the dynamic decision problem $P^{\text{coord}} = p_0^{\text{coord}}, P_1^{\text{coord}}, \ldots$ that is represented by the online decision model $M^{\text{coord}} = M_0^{\text{coord}}, M_1^{\text{coord}}, \ldots$.

Let $0 \leq f_{t_i} \leq 1$ quantify the portion of $R^+(t_i)$ that is copied into $R^+(t_i, f_{t_i})$. We define the current process punctuality rate (representing the current reliability of the transport system) by $p_t := \frac{l_t^{\text{punc}}}{l_t}$. There are l_t requests whose completion times (already realized or scheduled) fall into the period $[t - 500, t + 500]$ (moving time window). From these requests, l_t^{punc} requests have been completed or are scheduled to be completed within the time windows agreed with the customers (Schönberger and Kopfer 2009b).

In case that the current punctuality rate p_{t_i} is higher than the intended threshold rate p^{target}, f_{t_i} should be close to zero or even zero and $R^+(t_i, f_{t_i})$ should remain empty, which means no option will be exercised. If p_{t_i} has fallen below p^{target}, then f_{t_i} should be close to one or even equal to one, so that "$R^+(t_i, f_{t_i}) \approx R^+(t_i)$": for all additional requests an LSP-option is exercised. If p_{t_i} increases (decreases), then f_{t_i} decreases (increases) proportionally (step B in Fig. 1).

In the request choice new requests which are not compatible with the existing vehicle routes should be identified and preferentially put into $R^+(t_i, f_{t_i})$. Here, a new request r is compatible with the routes if r can be served so that "only slight modifications of the routes become necessary". Therefore, the coordinator agent needs to identify all new requests that cannot be served by the existing routes because

– the associated pickup and/or delivery location require detours and/or
– their associated pickup and/or delivery time window closes before a vehicle reaches the pickup or delivery site.

The marginal costs of a certain request can hardly be calculated so that substitute measures for the profitability/compatibility of a request r must be exploited instead. Those requests that seem to be least profitable or least compatible with the already scheduled requests are selected accordingly and put into the set $R^+(t_i, f_{t_i})$.

3.2 Dispatching task of the fleet-managing agent

In the investigated scenario, a customer site requires the visit of a vehicle (like in vehicle routing problems) within a time window. Such requests occur in applications where service team crews are sent to customer sites in order to remedy failures or perform some maintenance or repair work. Other application fields comprise the delivery of small standardized items stored in a large quantity on board of the vehicles and the collection of waste by call of the customer. In all applications, a time window is agreed in order to synchronize the customer and service team availability at the customer sites.

A single request r consecutively attains different states. Initially, r is known but it is not yet scheduled (K). Then, r is assigned to an own vehicle (I, short for internal fulfillment) or subcontracted (E, short for externalization). If the operation at the corresponding customer site has already been started but not yet been finished the state S (short for started request) is assigned to r. Requests completed after the last transportation plan update at time t_{i-1} are stored in the set $R^C(t_{i-1}, t_i)$. The new request stock $R(t_i)$ is determined by $R(t_i) := R(t_{i-1}) \cup R^+(t_i) \backslash R^C(t_{i-1}, t_i)$. Each request belongs at each time to exactly one of the sets $R^K(t_i)$, $R^E(t_i)$, $R^I(t_i)$ or $R^S(t_i)$, in which the requests having a common state are collected.

The problem of updating a transportation plan at time t_i is as follows. Let V denote the set of all own vehicles, $P_v(t_i)$ the set of all paths (sequence of visiting sites beginning with the position of the vehicle at time t_i and ending with the central depot) executable by vehicle v, and let $P(t_i)$ denote the union of the sets $P_v(t_i)$ $(v \in V)$. If the request r is served in a path p, then the binary parameter a_{rp} is set to 1; otherwise it is set to 0. A request r, already known at time t_{i-1} that is not subcontracted in TP_{i-1} is served by vehicle v_r. The travel costs associated with path p are denoted as $C^1(p)$. Finally, $C^2(r)$ gives the costs of exercising the LSP-option for request r.

In order to code the necessary decisions for determining a transportation plan, we deploy two families of binary decision variables. Let $x_{pv} = 1$ if and only if path $p \in P(t_i)$ is selected for vehicle $v \in V$ and let $y_r = 1$ if and only if request r is subcontracted.

$$\sum_{p \in P(t_i)} \sum_{v \in \mathcal{V}} C^1(p) x_{pv} + \sum_{r \in R(t_i)} C^2(r) y_r \to \min \tag{1}$$

$$\sum_{p \in P_v(t_i)} x_{pv} = 1 \quad \forall v \in \mathcal{V} \tag{2}$$

$$x_{pv} = 0 \quad \forall v \in \mathcal{V}, p \notin P_v(t_i) \tag{3}$$

$$y_r + \sum_{p \in P(t_i)} \sum_{v \in \mathcal{V}} a_{rp} x_{pv} = 1 \quad \forall r \in R(t_i) \tag{4}$$

$$y_r = 1 \quad \forall r \in R^E(t_i) \cup R(t_i, f_{t_i}) \tag{5}$$

$$\sum_{p \in P_{v(r)}} a_{rp} x_{pv_r} = 1 \quad \forall r \in R^S(t_i) \tag{6}$$

$$y_r \in \{0, 1\} \quad \forall r \in R(t_i), \;\; x_{pv} \in \{0, 1\} \quad \forall p \in P(t_i), \;\; v \in \mathcal{V}. \tag{7}$$

The costs for the updated transportation plan are minimized (1). A (possibly empty) path is assigned to each SC-vehicle (2) so that the selected path is executable by the selected vehicle (3). Each incomplete request must be scheduled so that either an LSP-option is exercised for it, or it is fulfilled by an SC-owned vehicle (4). Requests already externalized in a previous planning step or for which an LSP-option has been decided by the coordinator agent are assigned to an LSP (5). If the (un)loading operation of a request at a pickup (delivery) location has already been started, then this request cannot be re-assigned to another transport resource (6).

Originally, the online-formulation (1)–(4), (6), and (7) of the dynamic deployment problem has been proposed in (Hiller et al. 2006; Krumke et al. 2002;

Grötschel et al. 2002). To ensure that the coordinator decisions about the exercise of LSP-options are considered by the fleet managing agent during the re-deployment planning, the constraint (5) has been added (Schönberger and Kopfer 2007). This constraint is adjusted to the coordinator's decisions for every new instance of the update model.

The dynamic and deterministic model (1)–(7) is denoted as M_i^{fm}. It represents the instance P_i^{fm} of the online decision problem $P^{\mathrm{fm}} := P_0^{fm}, P_1^{\mathrm{fm}}, \ldots$ of the fleet managing (fm) agent.

An alternative modeling approach to integrate the solution of P_i^{coord} into M_i^{fm} is to remove all requests from the current request portfolio $R(t_i)$ for which the LSP-option has been chosen. The set $R(t_i)$ would then be replaced by the portfolio $R(t_i)\backslash(R^E(t_i) \cup R^+(t_i, f_{t_i}))$ and the constraint (5) could be deleted from the model (1)–(6). We have selected the "adaptive constraint"-representation of the deployment, because this formulation is straightforward and the adaptivity is striking in (5).

Previous investigations have revealed that a load peak can easily be managed if the subcontracting options have comparable costs. In this case the fleet manager, who wants to maximize his profit, selects an LSP-option in order to avoid late-request completions (and therefore prevents penalty payments). However, if the costs for selecting and using an LSP-option to fulfill request r dominate the costs associated with the fulfillment of r by a fleet management agent-controlled vehicle, then a profit-oriented transport service agent desists to draw an LSP-option. A load peak then causes a significant decrease of the process quality (lateness increase, etc.) (Schönberger and Kopfer 2009b). In the remainder of this article, we restrict our investigation to scenarios where the costs for exercising an option for a request r are significantly higher than the costs of the fleet management agent to use the own fleet to complete r.

With the now clearly described planning scenario in mind, we can concretize the research hypothesis stated in the introduction. *If the coordinator adaptively draws LSP-options (by fixing the control factor f_{t_i} with respect to p_{t_i}) then the punctuality rate p_{t_i} is increased during and after a peak situation compared to the case where the coordinator does not draw options ($f_{t_i} = 0$ for all t_i).*

4 Adaptive online vehicle dispatching

This section describes the main components of the developed dispatching system supporting the decision making in the previously introduced dynamic decision problem. In Subsect. 4.1, we introduce the used rolling horizon planning framework. The framework procedure for the transportation plan update exploring decision model adjustment is presented in Subsect. 4.2. We report a decision model pre-processing approach in Subsect. 4.3. The solving of the fleet managing agent's deployment problem is subject of Subsect. 4.4.

4.1 Rolling horizon planning framework

The basic idea of rolling horizon planning is to generate a sequence of plans $S_0, S_1, \ldots$ (S_i is a solution of the decision model M_i of instance P_i). At time t_i the plan S_i is derived and its realization is initiated. We continue with the execution of S_i until, at time t_{i+1}, additional data are revealed. The plan S_i is replaced by S_{i+1} and S_{i+1} is executed until it is corrupted by additional data and so on. The solution update is carried out if a pre-specified time point is reached (Rajagopalan et al. 2008; Saéz et al. 2008; Erera et al. 2008; Mitrović-Minić and Krishnamurti 2004), or if one or more certain events take place (Ausiello et al. 1994; Lund et al. 1996; Fleischmann et al. 2004).

Two general concepts for the update of S_i to S_{i+1} are distinguished. *Rule-based* updating follows the hypothesis that a few basic reasoning rules are valid and that it is possible to inductively reason the behavior in all other cases not explicitly stated in the basic rules (Lindstaedt 2007). The a-priori-route-concept (Tang and Miller-Hooks 2007; Liu 2007) is an example of rule-based updating. Additionally, update rules like MST-algorithms (Ausiello et al. 1994) or cheapest insertion approaches (Fleischmann et al. 2004) are representative examples for rule-based reasoning. A deductive reasoning is carried out in *model-based* update. Here, the set of all possible update alternatives is implicitly described by a formalized problem description (the decision model) and a structured scanning of the set of alternatives leads to the desired solution. Examples of model-based approaches include the linear programming-based optimization of a traveling salesman's route and the solving of the capacitated vehicle routing problem. Dynamic pickup and delivery problems with model-based schedule update techniques are investigated (among others) in Lund et al. (1996), Mitrović-Minić and Krishnamurti (2004), and Saéz et al. (2008). A necessary prerequisite is the availability of a suitable decision model. For the problem introduced in Sect. 3, the online optimization model with the instance representation (1)–(7) prepares a model-based solving in a rolling horizon planning framework.

Re-planning approaches that only consider the currently and surely known planning data without incorporating expected data are called *myopic*. They are based on the assumption that each forecast is wrong because the future events that will corrupt the execution of once generated processes are unpredictable (Fleischmann et al. 2004). In contrast, re-planning approaches with *anticipation* exploit forecasts (Powell 1996; Saéz et al. 2008) in which process decisions are likely to be executed as planned. In this article, we propose a special myopic re-planning approach. This approach is motivated by the impossibility to forecast future process disturbing events.

4.2 Rolling horizon planning with model preprocessing

The optimization model (1)–(7) represents the planning task of the subordinate transport service agent. It is the core of a rolling horizon planning system. The coordinator decision about the exercise of LSP-options for the requests $r \in R(t_i, f_{t_i})$ are reflected into the model of the current decision instance by fixing $y_r = 1$ for all $r \in R(t_i, f_{t_i})$. Thus, a preselection of values for some decision variables is carried out before the solving of the next instance of the online model M_i^{fm}. Techniques to preselect values of decision variables are subsumed under the term *presolving* (Andersen and Andersen 1995). Presolving is normally used to erase redundancies in a proposed decision model without shrinking or even extending the set of feasible model solutions. However, we can also use this technique to intervene in the online decision process by fixing the values of selected variables of the model (1)–(7). Doing so, we shrink the set of feasible solutions that is scanned by the applied search procedure.

In the context of principal–agent relationships, we can adjust the decision space of the subordinate agent if the principal agent applies presolving. Actually, we equip the principal to analyze the model data and to manipulate the domain of the decision variables $y_r, r \in R(t_i, f_{t_i})$ before the data are used to define the instance M_i^{fm}. Such a data analysis and model manipulation are referred to as *decision model preprocessing* (Solnon 2002). Here, model preprocessing is applied if the process punctuality p_{t_i} falls below the given threshold p^{target} or if it runs into danger to fall below p^{target} after the next additional requests have been released. By fixing $y_r = 1$ for the requests $r \in R(t_i, f_{t_i}) \subseteq R^+(t_i)$, the coordinator extends the LSP-usage in order to ensure or in order to re-increase the reliability of the processes.

Bierwirth (1999) discusses the controlled manipulation of a formal process control problem under the term *image modification* and proposes a generic closed-loop model control system in which the instances of an online optimization model are subject of control. Image modification represents an opportunity to synchronize an evolving decision situation with its formal representation (model), because it enables an effective and automatic context-based adjustment of the model.

Although the synchronization of a formalized problem situation with the motivating real-world situation has been evaluated positively in the control of computer systems (Arnold et al. 2005; Šegvić et al. 2006) the precise adjustment of mathematical optimization models for the control of larger-scale logistic systems has received only minor attention so far. In the remainder of this article we propose an implementation of the idea to couple the definition of the decision model M_i^{fm} of the current instance P_i with the current performance of the generated processes. The coupling enables a consideration of current data and a contextual model re-formulation.

More concretely, the decision model (1)–(7) is altered taking into account the current value of an error signal $e(p_{t_i})$, which expresses the current deviation of the punctuality p_{t_i} from the desired threshold p^{target}. The model (1)–(7) is an optimization model. It is possible to alter the objective function (1) or the constraint set (2)–(7) of the last instance M_{i-1}^{fm} during the definition of the new instance M_i^{fm}.

Gutenschwager (2002) proposes a first decoupling approach. He defines a local (temporary) planning objective for each single update instance at the re-planning time t_i. The applicability and suitability have been shown even for real-world applications (Gutenschwager and Böse 2003; Gutenschwager et al. 2004).

The planning framework for supporting the decision making in the situation outlined in Sect. 3 uses presolving and is as follows. Each of the two agents has to solve a decision problem in a re-planning cycle of the rolling horizon planning system. The superior coordinator agent has to decide for which requests the LSP-option is executed. Respecting these coordinator decisions, the subordinate transport service agent has to derive process updates.

A pseudo-code of the framework used to integrate the decision making of coordinator agent and the fleet managing agent is shown in Fig. 2. The process management procedure is controlled by two parameters: the *request selection strategy* Φ specifies the requests to be selected from $R^+(t_i)$ and to be put into $R(t_i, f_{t_i})$. A maximal

PROCEDURE process_management(Φ,β) ;
(a) i:=0;
(b) t_i:=GET_CURRENT_TIME();
(c) CurrentSolution:=GENERATE_INITIAL_SOLUTION();
(d) BROADCAST(CurrentSolution);
(e) **wait until** (CurrentSolution is completed) **or** (additional requests are released);
(f) **if** (CurrentSolution is completed) **then goto** (r);
(g) i:=i+1;
(h) t_i:= GET_CURRENT_TIME();
(i) $R^+(t_i)$:=GET_RELEASED_REQUEST(t_i);
(j) **if not** (SOLUTION_CORRUPTED(CurrentSolution)) **then goto** (e);
(k) p_{t_i}:=GET_CURRENT_PUNCTUALITY(t_i);
(l) $e(p_{t_i})$:=GET_CURRENT_ERRORSIGNAL(p_{t_i});
(m) $h_\beta(p_{t_i})$:=GET_INTERVENTION_INTENSITY($e(p_{t_i})$,β);
(n) $H(t_i, h_\beta(p_{t_i}))$:=SPECIFY_INTERVENTION($h_\beta(p_{t_i})$,$R^+(t_i)$,Φ);
(o) M_i^{fm}:=DEFINE_MODEL(t_i, CurrentSolution, $H(t_i, h_\beta(p_{t_i}))$);
(p) CurrentSolution :=SOLVE_MODEL(M_i^{fm}));
(q) BROADCAST(CurrentSolution);
(r) **Goto** (e);
(s) stop();

Fig. 2 Pseudo code of the algorithm framework

percentage β of additional requests to be selected from $R^+(t_i)$ is fixed in order to restrict the intervention of the coordinator agent (**maximal intervention intensity**).

Initially, the iteration counter i is set to 0 (a) and the first planning time is fetched (b). Next, an initial solution is generated, (c) and broadcasted to the vehicles of the transport service department and to the subcontractor(s) (d). Now, the procedure is idle and waits until the current solution has been completely executed or additional requests are received (e). In the first case, the procedure stops, (f) and is re-started as soon as additional processes are started. If the process execution is still in progress, then the iteration counter is increased by 1 (g) and the current system time t_i is fetched (h). All requests just released at time t_i are put into the set $R^+(t_i)$ (i). Next, it is checked whether the consideration of the additional requests compromises the current process (j). The procedure falls back into an idle state if no process corruption occurs. Otherwise, the current performance p_{t_i} (punctuality rate) is calculated (k), the error signal $e(p_{t_i})$ is derived (l), the intervention intensity $h_\beta(p_{t_i})$ is determined (m) and the requests which are prematurely directed into the SC fulfillment mode are selected (n). Afterwards, the new decision model is defined (o) and a high-quality solution of this model is derived (p) to replace the so far followed solution. The new solution is broadcasted to inform the field teams and the subcontractors (q). Again, the procedure falls back into the idle (waiting) state (r).

We have configured an adaptive process control system making use of image modification. Figure 3 shows the proposed configuration. Additional requests disturb the execution of the running processes at time t_i. As soon as this disturbance is detected an update of the currently executed transportation plan becomes necessary. Both, the current punctuality p_{t_i} and information about the disturbing additional requests in $R^+_{t_i}$ are forwarded to the model controller. The model controller consists of two consecutively processed stages. In the first stage, the intensity function h_β compares the current process punctuality p_{t_i} with the externally specified reference punctuality rate

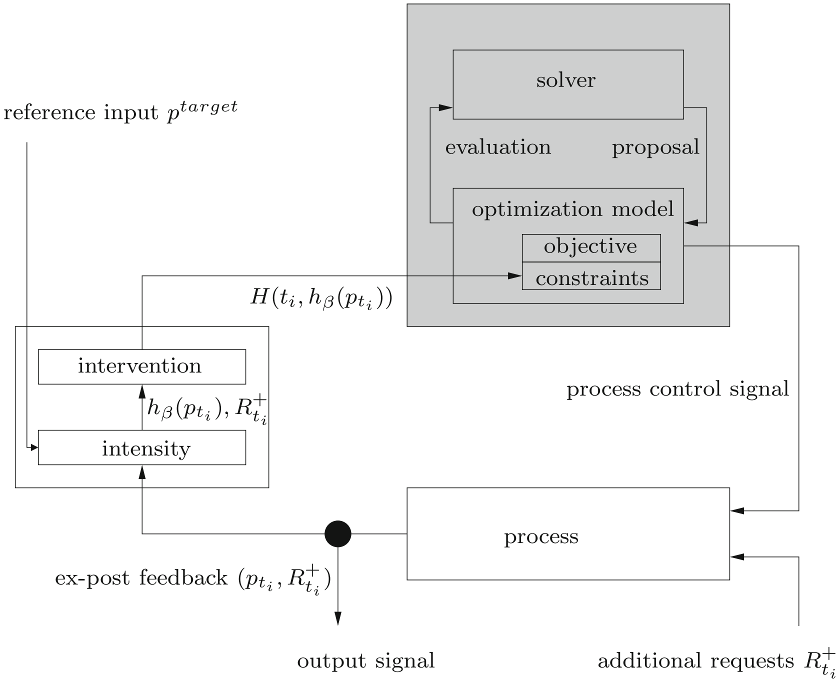

Fig. 3 Intervening an adaptive process control system with image modification

p^{target} [step (l) in Fig. 2]. The determined intervention intensity $h_\beta(p_{t_i})$ is forwarded to the intervention function H together with the set of additional requests [step (n) in Fig. 2]. Next, the intervention function derives an intervention signal $H(t_i, h_\beta(p_{t_i}))$ which is forwarded to the process controller (gray-shaded area in Fig. 3), which consists of a mathematical optimization model (defined by an objective function and a set of constraints) and a model solver. Instructions to adjust the constraint set to the current process feedback signal are found in the H-signal, which is directed to the constraint set. After the required adjustments have been established, the solver generates solution proposals which are evaluated using the updated decision model until the best proposal with regard to the current model instance is identified. This proposal is submitted as process control signal to the process and updates it. The updated process is executed until additionally arriving requests require a further process revision.

4.3 Solving the decision problem of the coordinator agent

We now introduce different strategies to determine the applied coordinator interventions. Basically, we propose four strategies to compile the set $R(t_i, h_\beta(p_{t_i}))$.

4.3.1 Reference strategies from literature

The idea of a *static strategy* is to let the coordinator make no intervention, e.g., fixing $f_{t_i} := 0$ and $R^+(t_i, f_{t_i}) := \emptyset$ for all re-planning times $t_1, t_2, \ldots$. Doing so, the fleet manager does not receive any adjustment instructions (step B in Fig. 1). Consequently, the fleet manager is solely responsible for the provision of a reasonable high-punctuality rate. Two static strategies have been proposed in order to ensure that a least punctuality rate is kept:

1. The fleet managing agent is obliged to meet the target punctuality rate p^{target} independently of the process costs. A strict and non-breakable planning restriction has to be respected (HARD).
2. Penalties are accounted to the fleet management agent's profit for each late arrival at a customer site (PEN).

The HARD-approach is not applicable because it shifts all risk arising from the planning uncertainty to the fleet managing agent (Schönberger and Kopfer 2009b). Furthermore, computational investigations and simulation experiments have shown that a fixed penalty as deployed in the PEN configuration is inappropriate to integrate the fulfillment of both dispatching goals "punctuality" (goal of the coordinator agent and the SC consortium) and "profitability" (goal of the fleet manager agent) (Schönberger and Kopfer 2007).

In an *adaptive strategy* the control parameter f_{t_i} is recurrently adjusted to the current value of the process punctuality rate p_{t_i}. The idea of Gutenschwager (2002) to temporarily substitute the global objective function is extended in Schönberger and Kopfer (2009b), where a new local objective function is defined for each new decision problem instance taking into account the current value of p_{t_i}. This local objective function is used to enable a context-related and -sensitive evaluation of the decisions made by the subordinate service agent. Since the variation of the objective function of an optimization procedure alters the search trajectory of an applied search algorithm, we refer to this strategy as *Search Direction ADaptation* (*SDAD*).

By adjusting the evaluation system for service agent decisions with SDAD, the coordinator agent can only indirectly intervene into the dispatching decisions of the fleet managing agent. The achievement of the intended alternation of the dispatching process cannot be guaranteed. To remedy this deficiency, we propose to use the model preprocessing approach outlined in Subsect. 4.2. By fixing the value $y_r = 1$ for selected requests the coordinator ensures that its intervention will surely become effective. Since the pre-selection of subcontracted requests affects the constraint set of the next instance of the online optimization model, we call this approach *Constraint Set ADaptation* (*CSAD*).

4.3.2 Constraint set adaptation

We re-use the control circuit for the rule-based determination of the coordinator interventions proposed for SDAD (Schönberger and Kopfer 2009b). In this control circuit, the current reference threshold p^{target} is compared to the current process punctuality p_{t_i} fetched by the function GET_CURRENT_PUNCTUALITY(t_i). The reference input is defined by $r(t_i) := [p^{\text{target}}; 1]$. This leads to the system development corridor $D(t_i) := [t_i; \infty) \times [p^{\text{target}}; 1]$ describing the desired future system performance and its core $C(t_i) := [t_i; \infty) \times [p^{\text{target}} + 0.1; 1]$ representing all future system states that do not require any model adjustment. We set $p^{\text{target}} := 0.8$, get the system development corridor $[t_i;\infty)\times[0.8;1]$ and its core $[t_i;\infty)\times[0.9;1]$. As long as $p_{t_i}\geq 0.9$ the current system performance (t_i, p_{t_i}) belongs to the core $C(t_i)$ and the system is safe in an HQ-state. If p_{t_i} falls below 0.9 and if the distance of p_{t_i} from 0.9 increases, then the system performance gets more and more off the core $C(t_i)$ and finally leaves even the system development corridor $D(t_i)$. This leads to the following error signal (8) that is calculated by calling the function: GET_CURRENT_ERRORSIGNAL(p_{t_i}).

$$e(p_{t_i}) := -\min(p_{t_i} - (p^{\text{target}} + 0.1); 0) \tag{8}$$

This error signal prematurely indicates if the system performance runs into danger of leaving the system development corridor as soon as the next external disturbance like a peak in the system workload occurs.

A controller transforms the previously calculated error signal $e(p_{t_i})$ into a control value used to determine the necessary manipulations of the existing decision model afterwards. Therefore, it is a mapping h_β that assigns the control value $h_\beta(p_{t_i})$ to the current process punctuality rate p_{t_i}. We define h_β as the piecewise linear function (9) which is calculated by calling GET_INTERVENTION_INTENSITY($e(p_{t_i})$) (Schönberger and Kopfer 2007).

$$h_\beta(p_{t_i}) = \begin{cases} 0, & e(p_{t_i}) \leq 0 \\ \beta, & e(p_{t_i}) \geq 0.2 \\ 5\cdot e(p_{t_i})\cdot\beta, & \text{in all other cases} \end{cases} \tag{9}$$

The value $h_\beta(e(p_{t_i}))$ is interpreted as percentage of the recently released requests at time t_i for which the SC-mode is selected by the coordinator agent.

The number $N_{t_i}^{\text{PRE}}$ of affected requests is determined as specified in (10). At time t_i an LSP-option is exercised for at most $\lceil\beta\cdot \mid R^+(t_i)\mid\rceil$ additionally arrived requests.

$$N_{t_i}^{\text{PRE}} := \lceil \mid R^+(t_i)\mid \cdot h_\beta(p_{t_i})\rceil. \tag{10}$$

No request is enforced into the SC-mode if the error signal is 0. The percentage of enforced externalization increases (decreases) smoothly and proportionally with an increasing (decreasing) error signal.

Finally, [corresponding to step (n) in the framework procedure in Fig. 2], the specification of the intervention is carried out by calling the function SPECIFY_INTERVENTION($h_\beta(p_{t_i})$;$R^+(t_i)$;Φ). The set $R(t_i, h_\beta(p_{t_i}))$ is returned by this function.

A sequencing rule Φ determines the order SEQ($R^+(t_i),\Phi$) of the elements in $R^+(t_i)$. It assigns a numerical value σ_r to each request $r\in R^+(t_i)$ and specifies whether $R^+(t_i)$ is ordered by increasing or by decreasing σ_r-values.

We first arrange the $m(i)$ elements contained in the set $R^+(t_i)$ of recently released requests in the sequence SEQ$(R^+(t_i),\Phi) := (r_{i_1}, r_{i_2},\ldots, r_{i_{m(i)}})$ according to Φ. Next, we initialize $R(t_i, h_\beta(p_{t_i})) := \emptyset$. Then, we consecutively insert the requests $r_{i_1}, r_{i_2},\ldots$ into the set $R(t_i, h_\beta(p_{t_i}))$. If the number of elements in $R(t_i, h_\beta(p_{t_i}))$ has reached the number $N_{t_i}^{\text{PRE}}$ we stop with the insertion of requests into $R(t_i, h_\beta(p_{t_i}))$ and set $H(t_i, h_\beta(p_{t_i}))$:= "Set $y_r = 1$ for all $r\in R(t_i, h_\beta(p_{t_i}))$". The call of the function DEFINE_MODEL(t_i, CurrentSolution, $R(t_i, h_\beta(p_{t_i}))$) triggers the formulation of the next decision model instance $M(t_i)$. The fulfillment mode of the remaining requests $R^+(t_i)\backslash R(t_i, h_\beta(p_{t_i}))$ can be freely determined by the fleet management agent.

We use the sequencing rule *isolation based sequencing* (*IBS*) introduced in Schönberger and Kopfer (2009). Expenses and benefits of a single request can hardly be evaluated since the coupling effects of combining the fulfillment of several requests are very high. This observation motivates the development and implementation of a rule that tries to identify those requests which cannot be combined efficiently with other requests. For these isolated requests LSP-options are preferentially exercised. In order to evaluate the "degree of isolation" of the site of a request $r \in R(t_i)$, we first calculate for each request r its distance $d_1(r)$ from the median $\text{med}(R^+(t_i))$ of the locations of requests contained in the current request portfolio. After having calculated this distance for each request in $R^+(t_i)$, we calculate the normalized distance $d_1^*(r) := \frac{d_1(r)}{\max\{d_1(r)|r\in R^+(t_i)\}}$ for each request $r \in R^+(t_i)$. If $d_1^*(r)$ is close to 1, then r is situated at the edge of the operations area which is often a first hint for isolation. To find out whether r can be combined with other requests into an efficient route, we calculate the distance $\text{mindist}(r)$ to the nearest other request site in the complete request portfolio $R(t_i)\backslash R^E(t_i)$, that has not yet been subcontracted. It is $\text{mindist}(r) := \min\{d_2(r, r_j) + d_3^{\text{tw}}(r, r_j)|r_j \in R(t_i)\backslash R^C(t_i)\}$, where $d_2(r, r_j)$ gives the travel distance between μ_r and μ_{r_j}. The term $d_3^{\text{tw}}(r, r_j)$ is used to depreciate the spatial distance in case that the time windows $\text{TW}(r) := [t_r^+, t_r^-]$ and $\text{TW}(r_j) = [t_{r_j}^+, t_{r_j}^-]$ of r and r_j interdict the combination of the two requests in one route. It is $d_3^{\text{tw}}(r, r_j) := 0$, if $\min\{|\ t_r^+ - t_{r_j}^-\ |, |\ t_j^+ - t_r^-\ |\} \geq \text{dist}(r, r_j)$ (that is, there is enough time for a vehicle to travel from μ_r to μ_{r_j} or vice versa) and in all other cases it is $d_3^{\text{tw}}(r, r_j) := \text{dist}(r, r_j) - \min\{|\ t_r + -t_{r_j}^-\ |, |\ t_{r_j}^+ - t_r^-\ |\}$. Finally, we calculate the normalized minimal distance indicator $\text{mindist}^*(r) := \frac{\text{mindist}(r)}{\max\{\text{mindist}(r)|r\in R^+(t_i)\backslash R^C(t_i)\}}$. The value (11) is then assigned as sorting value σ_r to request r. If σ_r is small (close to 0), then the site μ_r is either in the center of the operations area, or it is close to the sites of other requests. If a request site μ_r is situated at the edge of the operations area and not closely situated to the sites of other requests, then r can be classified as isolated ($\sigma_r \approx 1$). We now sort the requests in $R^+(t_i)$ by decreasing σ_r-values. At the beginning of the sequence of requests the most isolated requests are found and LSP-options for these requests are exercised first.

$$\sigma_r := d_1^*(r) \cdot \text{mindist}^*(r) \tag{11}$$

4.4 Solving the model of the fleet manager agent decision task

The adjusted model (1)–(7) is a very complex optimization model. We do not expect to solve it optimally and therefore propose a heuristic search approach to derive a solution of high quality. Besides several other meta-strategies (Rajagopalan et al. 2008; McKendall and Jaramillo 2006; Mitrović-Minić and Krishnamurti 2004) evolutionary approaches have proven their applicability for rolling horizon planning approaches to solve instances of an online decision problem (Saéz et al. 2008; Branke 2001).

For solving the instances of the online decision model we use a Memetic Algorithm realizing a hybrid search strategy consisting of a genetic search and a local 2-opt improvement procedure. Every time a new decision problem instance has been stated, the Memetic Search Algorithm is re-started by the call of the SOLVE_MODEL command [step (p) in the procedure in Fig. 2] (Schönberger 2005). In order to protect the decisions made by the coordinator, neither the application of cross-over nor mutation operators alter a coordinator decision for exercising an LSP-option.

5 Computational simulation experiments

After having introduced and configured adaptive coordinator interventions into the online optimization framework, we can re-formulate the motivating research hypothesis as follows: *If we use CSAD to determine f_{t_i} (and therefore fix the sets $R(t_i, f_{t_i})$, $i = 0, 1, 2, \ldots$) then the punctuality rate p_{t_i} is increased compared to the case where f_{t_i} is fixed to 0*. In this section, we report about simulation experiments carried out to prove this hypothesis. At first, we outline the simulated scenarios (Subsect. 5.1). Next, we describe the performance indicators observed during the simulations in order to assess the different strategies (Subsect. 5.2). Finally, we describe and discuss the observed performances for the different strategies (Subsect. 5.3).

5.1 Simulated scenarios

A *scenario* $(\text{exp}, \omega, P, \alpha)$ is determined by applying a *planning system setting* (exp, ω) to an *incoming stream of requests* (P, α).

We use four streams of incoming requests during the simulation period [1000:5000]. In each stream the requests are randomly drawn from one of the Solomon instances $P \in \mathcal{P} := \{R103, R104, R107, R108\}$ (Solomon 1987). Each stream is balanced, but a demand peak occurs between time 1500 and 1700. As mentioned at the end of Subsect. 3.2 the process planning is challenging if and only if the costs for drawing a subcontraction option are significantly higher than the costs for the self-fulfillment by the own fleet. Therefore, we restrict our computational experiments and simulate only scenarios where $\alpha = 3$ (the exercise of an LSP-option costs three times the expenditures of the own vehicle usage including necessary penalty payments).

The determination of the system setting requires the specification of the intervention strategy $\text{exp} \in \mathcal{S} := \{\text{HARD, PEN, SDAD, CSAD}\}$ used by the supply net coordinator and the seeding $\omega \in \mathcal{O} := \{1, 2, 3\}$ of the MA used by the transport service agent. If we apply CSAD, then we use IBS to select those requests for which the LSP-option is pre-selected. Previous experiments have revealed that the efficiency of CSAD increases if the maximal intervention intensity β is lifted. For this reason, we restrict our simulation experiments to planning system settings with $\beta = 1$ (Schönberger and Kopfer 2007). Overall, we investigate $\mid \mathcal{S} \mid \cdot \mid \mathcal{O} \mid = 4 \cdot 3 = 12$ planning system settings.

Combining the previously mentioned parameter settings, we setup $12 \cdot \mid \mathcal{P} \mid = 12 \cdot 4 = 48$ scenarios for simulation. In each scenario, a target punctuality $p^{\text{target}} = 0.8$ must be met.

5.2 Performance indicators

We define two groups of process quality indicators. The indicators in the first group measure the non-monetary performance (punctuality, backlog size, etc.), and the indicators in the second group provide insights into the costs of the generated transportation plans.

Process quality indicators. The punctuality recorded at time t within the scenario $(P, \exp, \omega)$ is denoted as $p_t(P, \exp, \omega)$. In all scenarios, we have $\alpha = 3$, so that we skip α from the scenario description. Let $p_t(\exp) := \frac{1}{12}\sum_{\omega\in\mathcal{O}}\sum_{P\in\mathcal{P}} p_t(P, \exp, \omega)$ denote the average punctuality observed at time t for the strategy exp.

In order to study the impacts of the demand peak with respect to the punctuality p_t, we calculate the deviation of $p_t(P, \exp, \omega)$ from the reference value $p_{1000}(P, \exp, \omega)$ for all times in the observation time interval $[1000, 5000]$ by $p_t(P, \exp, \omega)/p_{1000}(P, \exp, \omega) - 1$.

The largest past-peak deviation from the reference value is then calculated by $\frac{\min_{t\geq 1500}\{p_t(P,\exp,\omega)\}}{p_{1000}(P,\exp,\omega)} - 1$. Now, the average $\delta(\exp)$ of the largest observed deviation from the reference values for the strategy exp applied is given by

$$\delta(\exp) := \frac{1}{12}\sum_{\omega\in\mathcal{O}}\sum_{P\in\mathcal{P}}\left(\frac{\min_{t\geq 1500}\{p_t(P, \exp, \omega)\}}{p_{1000}(P, \exp, \omega)} - 1\right).$$

Let $T^{\text{below}}_{\exp}$ denote the first time in which $p_t(\exp)$ falls below p^{target} and $T^{\text{heal}}_{\exp} := \min\{t \in [1000, 5000] \mid \nexists\, l \in [t, 5000], p_l < p^{\text{target}}\}$ referring to the time in which an HQ state is finally re-achieved by $p_t(\exp)$. We define $(\exp) := \dfrac{T^{\text{heal}}_{\exp} - T^{\text{below}}_{\exp}}{4000}$ as the percentage of LQ periods within the observation interval $[1000, 5000]$.

Throughout the simulation time, we have recorded the percentage of subcontracted requests in $q_t(P, \exp, \omega)$. The average of these values has been calculated and stored in $q_t(\exp) := \frac{1}{12}\sum_{\omega\in\mathcal{O}}\sum_{P\in\mathcal{P}} q_t(P, \exp, \omega)$ for each strategy exp. The maximally observed subcontraction rate is stored in $\sigma(\exp) := \max_{t\geq 1500} q_t(\exp)$. It indicates the exploitation intensity of the subcontraction fulfillment mode.

In order to get information about requests still uncompleted at time t, we fetch the number $w_t(P, \exp, \omega)$ of already scheduled but not fulfilled requests (*pending requests*) during the simulation of scenario $(\exp, P, \omega)$. We calculate the averagely observed number $w_t(\exp)$ of requests being pending at time t.

Financial process evaluation indicators. The cumulated request fulfillment costs observed in a scenario at time t are saved in $c_t(P, \exp, \omega)$ and their averages $c_t(\exp) := \frac{1}{12}\sum_{\omega\in\mathcal{O}}\sum_{P\in\mathcal{P}} c_t(P, \exp, \omega)$ are calculated. We approximate the marginal costs $mc_t(\exp)$ of a request for each strategy exp at time t. To calculate $mc_t(\exp)$, we first determine the increase $c_t(\exp) - c_{t-100}(\exp)$ of the cumulated costs at time t compared to the last replanning time $t - 100$. Second, we determine the number of requests $\mid R^C_{(t-100,t)}(\exp) \mid$ completed in the period between time $t - 100$ and t. The marginal

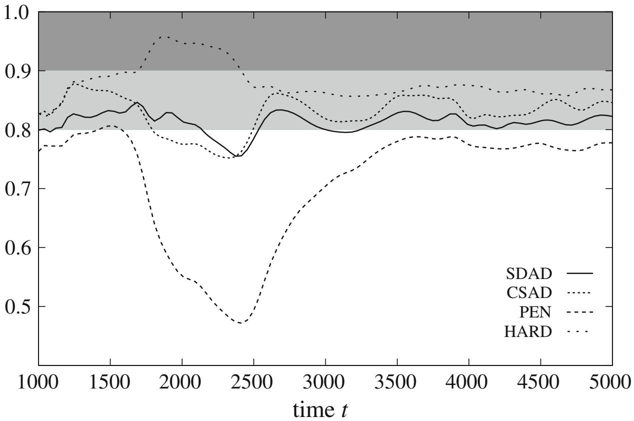

Fig. 4 Development of the punctuality $p_t(\text{exp})$

cost indicator $mc_t(\text{exp})$ is then defined as $mc_t(\text{exp}) := \frac{c_t(\text{exp})-c_{t-100}(\text{exp})}{|R^C_{(t-100,t)}(\text{exp})|}$. This value approximates the fulfillment costs for one single request at time t if exp is used. In order to compare the three alternative strategies PEN, SDAD, and CSAD with HARD (the reference approach), we calculate the increase $\gamma_t(\text{exp}) := \frac{c_t(\text{exp})}{c_t(\text{HARD})} - 1$ of the overall costs.

5.3 Presentation and discussion of results

In an online evaluation, we track the evolution of the performance indicators during a simulation of a scenario, and in an offline evaluation, we analyze the performance parameter after the completion of a simulation.

Online process quality evaluation. Figure 4 contains the punctuality rates $p_t(\text{exp})$ observed online during the simulation experiments. We observe that CSAD performs worse that HARD but significantly better than PEN. A comparable behavior of CSAD and SDAD is revealed. Similar to SDAD, the punctuality falls out of the system development corridor (the gray shaded area in Fig. 4) immediately after the load peak has started. SDAD is able to keep $p_t(\text{SDAD})$ within the corridor until time $t = 2200$ but $p_t(\text{CSAD})$ leaves the system development corridor already at time $t = 1900$. However, in both cases (CSAD and SDAD) the least observed punctuality is 0.75 and in both cases the system corridor is re-entered around time $t = 2500$. After the immediate reaction to the load peak is over, CSAD maintains as higher punctuality rate than SDAD. Furthermore, $p_t(\text{SDAD})$ leaves the system development corridor for a certain period around $t = 3000$ but $p_t(\text{CSAD})$ remains completely within the corridor for the remaining simulation time.

Since the punctuality $p_t(\text{CSAD})$ leaves the core of the system development corridor (the dark gray shaded area in Fig. 4) a non-zero error signal $e(t_i)$ defined by (8) is generated inducing a non-zero control signal (the intervention intensity) $h_1(p_{t_i})$

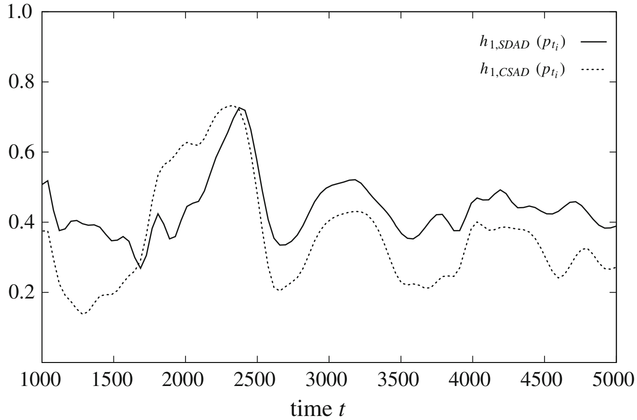

Fig. 5 Evolution of the intervention intensities

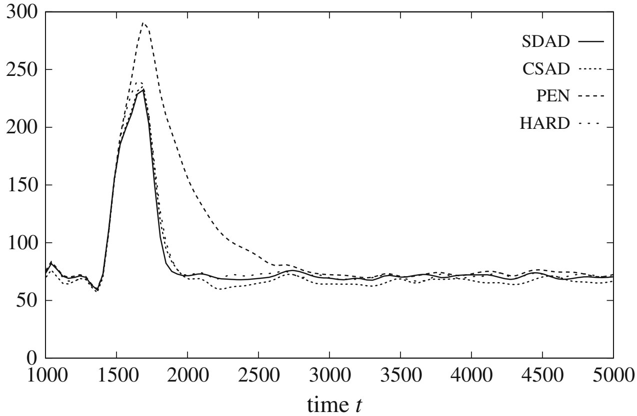

Fig. 6 Number of pending requests w_t (exp)

(9). The CSAD intervention intensity $h_{1,\mathrm{CSAD}}(p_{t_i})$ is compared with the SDAD intervention intensity $h_{1,\mathrm{SDAD}}(p_{t_i})$ in Fig. 5. In off-peak-periods, $h_{1,\mathrm{CSAD}}(p_{t_i}))$ is remarkably smaller than $h_{1,\mathrm{SDAD}}(p_{t_i}))$, but in an acute peak management period ($1700 \leq t \leq 2400$) the CSAD-intervention intensity is higher than the SDAD-intervention intensity.

The increase of the supply net coordinator intervention intensity immediately after the beginning of the load peak contributes to keep the number of scheduled, but not completed, requests (pending requests) in the system on a lower level than PEN is able to achieve (Fig. 6). Similar to HARD and CSAD, the number of pending requests does not climb above 240, and the averagely observed pre-peak number of pending requests is also re-achieved before $t = 2000$.

The analysis of the number of used own vehicles v_t (exp) and of the percentage q_t (exp) of externalized requests reveals the impacts of the higher intervention intensity of CSAD compared to SDAD. At first, the higher intervention intensity lifts up

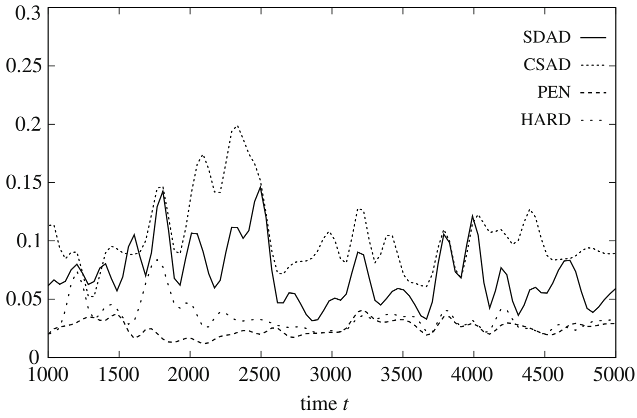

Fig. 7 Portion q_t (exp) of externalized requests in the schedule generated at time t

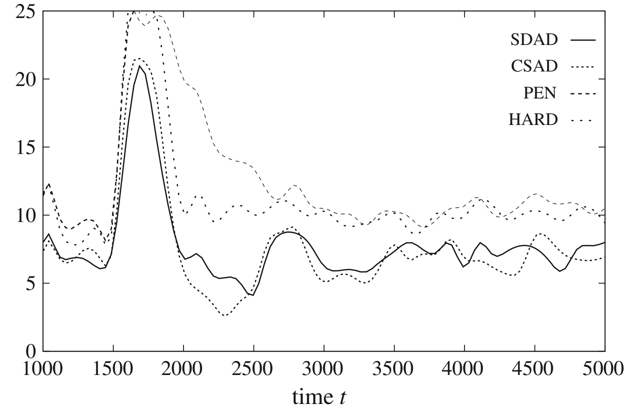

Fig. 8 Number of used own vehicles v_t (exp)

the post-peak number percentage of subcontracted requests (Fig. 7) during the interval from $t = 1900$ until $t = 2400$. In this period, the number of externalized requests is doubled compared to the SDAD case and quadruplicated compared to PEN and HARD. In the same period, fewer own vehicles are routed (Fig. 8). Compared to v_t (SDAD) the number v_t (CSAD) is reduced by 50% during the aforementioned period; compared with v_t (HARD) it is reduced by 66% and in comparison with v_t (PEN) it is reduced by 80%. In later stages of the simulation experiments, the numbers of used own vehicles v_t (SDAD) and v_t (CSAD) are nearly equal. However, the percentage q_t (SDAD) is significantly lower than q_t (CSAD).

Offline process quality evaluation. Table 1 contains all indicator values resulting from the offline evaluation of the planning system after the simulations have been finished. Again, we state that the performance of the adaptive strategies SDAD and CSAD clearly outperform PEN, but they are dominated by HARD. In detail, we first see that

Table 1 Offline process quality performance indicator values

	Exp			
	HARD (%)	PEN (%)	SDAD (%)	CSAD (%)
δ(exp)	3.5	−38.8	−5.6	−8.7
π(exp)	–	97.5	16.7	5.0
σ(exp)	8.0	4.1	14.5	19.7

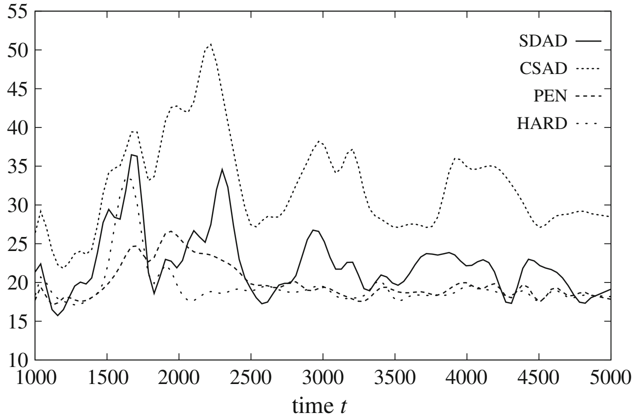

Fig. 9 Marginal costs mc_t(exp)

CSAD comes along with a slightly more severe maximal punctuality deviation than SDAD, e.g., $\delta(\text{SDAD}) = -5.6\% \geq -8.7\% = \delta(\text{CSAD})$. Second, we find out that less LQ-situations are achieved if CSAD is used (compared to SDAD): $\pi(\text{CSAD}) = 5.0\% \leq 16.7\% = \pi(\text{SDAD})$. Thus, CSAD produces more reliable sequences of transportation plans. Finally, we observe that CSAD is able to intervene more severely and draws more often an LSP option. It is $\sigma(\text{CSAD}) = 19.7\% \geq 14.5\% = \sigma(\text{SDAD})$. In conclusion, CSAD demonstrates a better performance with respect to reliability and intervention severity. However, it seems to be slightly outperformed by SDAD with respect to the maximal punctuality decrease after the load peak initiation.

Online-evaluation of the process costs. We have traced the average marginal costs of the completed requests throughout the simulations (Fig. 9). CSAD produces the highest marginal costs among all four analyzed approaches, and the largest oscillation amplitude is also observed for CSAD. The increased intervention intensity $h_{1,\text{CSAD}}(p_{t_i})$ during the past peak period $t \in [2000; 2400]$ leads to a significantly increased LSP option usage (compared the all three other strategies), which results in additional expenditures caused by the high-LSP tariffs. Conclusively stated, the two adaptive strategies CSAD and SDAD cause higher marginal costs than the non-adaptive strategies HARD and PEN. In addition, CSAD is the most cost-intensive strategy among the four analyzed approaches.

Table 2 Offline process costs performance indicator values

	Exp			
	HARD	PEN	SDAD	CSAD
$c_{5000}(\text{exp})$	56301.5	55748.3	64225.6	89599.4
$\gamma_{5000}(\text{exp})$	–	-1.0%	14.1%	59.1%

Offline process cost assessment. Table 2 contains the cumulated average process execution costs observed after the completion of the simulation runs. For both non-adaptive strategies PEN and HARD, we observe nearly the same costs $c_{5000}(\text{PEN})$ and $c_{5000}(\text{HARD})$. We see that the costs from the PEN-experiments are decreased by $\gamma_{5000}(\text{PEN}) = 1\%$ compared to the results achieved in the HARD-experiments. The two adaptive strategies produce significantly higher costs: $c_{5000}(\text{SDAD}) = 64225.6$ ($\gamma_{5000}(\text{SDAD}) = 14.1\%$) and $c_{5000}(\text{CSAD}) = 89599.4$ ($\gamma_{5000}(\text{CSAD}) = 59.1\%$).

6 Conclusions

We have analyzed a dynamic transport process planning problem that represents a typical decision scenario in supply chains. In order to enable a superior coordinator to intervene into the deployment decisions of a subordinate fleet manager, we have proposed an extended rolling horizon planning framework for online dispatching. Using model pre-processing selected decision variables are fixed so that requirements of the superior supply chain coordinator are surely respected by the subordinate transport service providing agents in its resource allocation.

Within computational simulation experiments, we observed the behavior of the processes in the investigated transport system. Several configurations of the process planning system have been set up. Results observed in the different configurations have been analyzed and compared. The initially formulated research hypothesis was verified.

The main contribution to the development and understanding of the control of complex logistic systems is that interventions of a superior supply chain coordinator are reasonable. We have proposed a contribution to the increase of the performance of a transport system by externalizing incompatible requests so that the system's own resources are exonerated from additional load. The developed prototype proves that an adaptive coupling of the planning systems of a superior coordinator and a subordinate service providing agent is possible. The managerial impacts of this finding must be further investigated, but we think that an adaptive integration of planning tools will result in additional improvements of the planning and management of supply chain processes.

Future research will follow two directions. On the one hand, we will concentrate on the prevention of the entrance of unprofitable requests into the system. The idea of using adaptive decision models seems to be promising also for capacity control purposes. On the other hand, we will transfer the acquired experience about the adjustment of decision models into the application field of horizontal cooperations of freight carrier companies. Here, interconnected optimization problems (Krajewska et al. 2008) require a regular reformulation in order to achieve a fair load balance between the autonomous partners.

Acknowledgments This research was supported by the German Research Foundation (DFG) as part of the Collaborative Research Center 637 "Autonomous Cooperating Logistic Processes" (Subproject B7).

References

Albers S, Charikar M, Mitzenmacher M (2001) Delayed information and action in on-line algorithms. Inf Comput 170(2):135–152
Andersen ED, Andersen KD (1995) Presolving in linear programming. Math Program 71:221–245
Arnold M, Fink SJ, Grove D, Hind M, Sweeney PF (2005) A survey of adaptive optimization in virtual machines. In: Proceedings of the IEEE 2005 93(2):449–466
Ausiello G, Feuerstein E, Leonardi S, Stougie L, Talamo M (1994) Serving requests with on-line routing. In: Schmidt EM, Skyum S (eds) Proceedings of 4th Scandinavian Workshop on algorithm theory. Springer, Berlin, pp 37–48
Bierwirth C (1999) Adaptive search and the management of logistics systems. Kluwer, Dordrecht
Black F, Scholes M (1973) The pricing of options and corporate liabilities. J Political Econ 81(3):637–654
Brage Illa A (1966) Critical probabilities and determinism in decision theory. Calcolo 2(2):117–130
Branke J (2001) Evolutionary optimization in dynamic environments. Kluwer, Dordrecht
Branke J, Mattfeld DC (2005) Anticipation and flexibility in dynamic scheduling. Int J Prod Res 43(15):3103–3129
Brehmer B (1992) Dynamic decision making: human control of complex systems. Acta Psychologica 81:211–241
Busemeyer JR (2001) Dynamic decision making. In: International encyclopedia of the social and behavioral sciences: methodology. Mathematic and computer science. Elsevier, Amsterdam, pp 3903–3908
Calza F, Passaro R (1997) EDI network and logistics management at Unilever-Sagit. Supply Chain Manag Int J 2(4):158–170
Crainic T, Laporte G (1997) Planning models for freight transportation. Eur J Oper Res 122:409–438
Delaitre L, Awasthi A, Molet H, Breuil D (2007) A multiple container loading problem based algorithm for efficient allocation of goods to vehicles. In: Proceedings of the 18th IASTED International Conference. pp 313–317
Erera A, Karacık B, Savelsbergh M (2008) A dynamic driver management scheme for less-than-truckload carriers. Comput Oper Res 35:3397–3411
Feillet D, Dejax P, Gendreau M (2005) Traveling salesman problems with profits. Transp Sci 39:188–205
Fleischmann B, Gnutzmann S, Sandvoß E (2004) Dynamic vehicle routing based on online traffic information. Transp Sci 28:420–433
Fleischmann B, Gietz M, Gnutzmann S (2004) Time-varying travel times in vehicle routing. Transp Sci 38(2):160–173
Garfinkel RS (1985) Motivation and modeling. In: Lawler EL, Lenstra JK, Rinnooy Kan AHG, Shmoys DB (eds) The traveling salesman problem. Wiley, New York
Gendreau M, Laporte G, Semet F (2004) Heuristics and lower bounds for the bin packing problem with conflicts. Comput Oper Res 31(3):347–358
Golden BL, Assad AA (eds) (1988) Vehicle routing: methods and studies. Elsevier, North-Holland
Golden B, Raghavan S, Wasil E (eds) (2008) The vehicle routing problem. Springer, New York
Grötschel M, Krumke SO, Torres LM (2002) Online-dispatching of automobile service units. In: Leopold-Wildburger U, Rendl F, Wäscher G (eds) Operations research proceedings. Springer, Berlin, pp 168–173
Gutenschwager K (2002) Online-Dispositionsprobleme in der Lagerlogistik. Physica-Verlag, Heidelberg
Gutenschwager K, Böse J, Voß S (2003) Effiziente Prozesse im Kombinierten Verkehr—Ein neuer Lösungsansatz zur Disposition von Portalkränen. Logistikmanagement 5(1):62–73
Gutenschwager K, Niklaus C, Voß S (2004) Dispatching of an electric monorail system: applying metaheuristics to an online pickup and delivery problem. Transp Sci 38(4):434–446

Hadjiconstantinou E, Roberts D (2002) Routing under uncertainty: an application in the scheduling of field service engineers. In: Toth P, Vigo D (eds) The vehicle routing problem. SIAM, Philadelphia, pp 331–353

Hiller B, Krumke SO, Rambau J (2006) Reoptimization gaps versus model errors in online-dispatching of service units for ADAC. Discret Appl Math 154:1897–1907

Ichoua S, Gendreau M, Potvin J-Y (2003) Vehicle dispatching with time-dependent travel times. Eur J Oper Res 144:379–396

Jensen M, Meckling W (1976) Theory of the firm. Managerial behavior, agency costs, and ownership structure. J Financial Econ 3(4):305–360

Kalleberg AL, Reynolds J, Marsden PV (2003) Externalizing employment: flexible staffing arrangements in US organizations. Soc Sci Res 32:525–552

Kaluza B, Dullnig H, Malle F (2003) Principal-Agent-Probleme in der Supply Chain—Problemanalyse und Diskussion von Lösungsvorschlägen. Discussion paper 2003/03 of the College of Business Administration. University of Klagenfurt, Austria

Kopfer H, Krajewska M (2007) Approaches for modelling and solving the integrated transportation and forwarding problem. In: Corsten H, Missbauer H (eds) Produktions- und Logistikmanagement. Verlag Franz Vahlen, Munich pp 439–458

Krajewska M, Kopfer H, Laporte G, Ropke S, Zaccour G (2008) Horizontal cooperations among freight carriers: request allocation and profit sharing. J Oper Res Soc 59(11):1483–1491

Krajewska M, Kopfer H (2009) Transportation planning in freight forwarding companies—Tabu Search algorithm for the integrated operational transportation planning problem. Eur J Oper Res (to appear)

Krumke SO, Rambau J, Torres LM (2002) Realtime-dispatching of guided and unguided automobile service units with soft time windows. In: Möhring RH, Raman R (eds) Algorithms—ESA 2002, 10th Annual European Symposium, Lecture Notes in Computer Science 2461, Springer, Berlin

Larsen A, Madsen OGB, Solomon MM (2008) Recent Developments in dynamic vehicle routing. In: Golden B, Raghavan S, Wasil E (eds) The vehicle routing problem. Springer, New York, pp 199–218

Laux H (1982) Entscheidungstheorie. Springer, Heidelberg

Lindstaedt H (2007) Problemlösen und Verstehen bei ökonomischen Agenten—eine Gegenüberstellung ökonomischer und kognitionspsychologischer Modelle regelbasierten Entscheidens. NeuroPsychoEconomics 2:30–43

Littman ML, Dean TL, Kaelbling LP (1995) On the complexity of solving Markov decision problems. In: Proceedings of the 11th Conference on uncertainty in artificial intelligence. pp 394–402

Liu Y-H (2007) A hybrid scatter search for the probabilistic traveling salesman problem. Comput Oper Res 34:2949–2963

Lo S-C, Hall RW (2008) The design of real-time logistics information system for trucking industry. Comput Oper Res 35:3439–3451

Lund K, Madsen OBG, Rygaard JM (1996) Vehicle routing problems with varying degrees of dynamism. Institute of Mathematical Modelling (IMM), Technical University of Denmark, Lyngby, Technical Report, No. IMM-REP-1996-1

Madsen OGB, Tosti K, Vælds J (1995) A heuristic method for dispatching repair men. Ann Oper Res 61:213–226

McKendall AR Jr, Jaramillo JR (2006) A tabu search heuristic for the dynamic space allocation problem. Comput Oper Res 33:768–789

Mitrović-Minić S, Krishnamurti R, Laporte G (2004) Double-horizon based heuristic for the dynamic pickup and delivery problem with time windows. Transp Res B 38:635–655

Moreira LM, Oliveira JF, Gomes AM, Ferreira JS (2007) Heuristics for a dynamic rural postman problem. Comput Oper Res 34:3281–3294

Munera HA (1984) The generalized means model (GMM) for non-deterministic decision making: its normative and descriptive power, including sketch of the representation theorem. Theor Decis 18:173–202

Pankratz G (2002) Speditionelle Transportdisposition. DUV Gabler, Wiesbaden

Potvin J-Y, Xu Y, Benyahia I (2006) Vehicle routing and scheduling with dynamic travel times. Comput Oper Res 33 :1129–1137

Powell WB (1996) A stochastic formulation of the dynamic assignment problem, with an application to truckload motor carriers. Transp Sci 30(3):195–219

Powell WB, Cheung RK (2000) Adaptive labeling algorithms for the dynamic assignment problem. Transp Sci 34(1):50–66

Psaraftis HN (1988) Dynamic vehicle routing problems. In: Golden BL, Assad AA (eds) Vehicle routing: methods and studies. Elsevier, North-Holland

Psaraftis HN (1995) Dynamic vehicle routing: status and prospects. Ann Oper Res 61:143–164

Rajagopalan HK, Saydam C, Xiao J (2008) A multiperiod set covering location model for dynamic redeployment of ambulances. Comput Oper Res 35(3):814–826

Ramamritham K, Stankovic JA (1994) Scheduling and operating system support for real-time systems. In: Proceedings of the IEEE. pp 55–67

Rego C, Roucairol C (1995) Using Tabu search for solving a dynamic multi-terminal truck dispatching problem. Eur J Oper Res 83:411–429

Ronen B, Coman A, Schragenheim E (2001) Peak management. Int J Prod Res 39(14):3183–3193

Saéz D, Cortéz CE, Núñes A (2008) Hybrid adaptive predictive control for the multi-vehicle dynamic pick-up and delivery problem based on genetic algorithms and fuzzy clustering. Comput Oper Res 35(11):3412–3438

Schönberger J (2005) Operational freight carrier planning. Springer, Heidelberg

Schönberger J, Kopfer H (2007) On decision model adaptation in online optimization of a transport system. In: Günther H-O, Mattfeld DC, Suhl L (eds) Management logistischer Netzwerke. Physica-Verlag, Heidelberg, pp 361–381

Schönberger J, Kopfer H (2007) On the value of objective function adaptation in online optimization. In: Kalcsics J, Nickel S (eds) Operations research proceedings 2007. Springer, Berlin, pp 491–496

Schönberger J, Kopfer H (2009a) Transport system responsiveness improvement. Int J Phys Distrib Logist Manag 39(1):63–79

Schönberger J, Kopfer H (2009b) Online decision making and automatic decision model adaptation. Comput Oper Res 36(6):1740–1750

Schönberger J, Kopfer H (2008) Schedule nervousness reduction in transport re-planning. Commun SIWN 3:7–14

Schönberger J, Kopfer H (2009) Rules for the identification of portfolio-incompatible requests in dynamic vehicle routing. In: Hansen HR, Karagiannis D, Fill H-G (eds) Proceedings of Wirtschaftsinformatik 2009, Band 2

Séguin R, Potvin J-Y, Gendreau M, Crainic TG, Marcotte P (1997) Real-time decision problems: an operations research perspective. J Oper Res Soc 48(2):162–174

Šegvić S, Remazeilles A, Chaumette F (2006) Enhancing the point feature tracker by adaptive modelling of the feature support. In: Leonardis A, Bischof H, Pinz A (eds) ECCV 2006. Springer, Berlin, pp 112–124

Seiden S (1996) Randomized online scheduling with delivery times. J Comb Optim 3:399–416

Sellmaier S (2007) Langfristiges Entscheiden. LIT Verlag, Berlin-Hamburg-Münster

Solnon C (2002) Boosting ACO with a preprocessing step. In: Cagnoni S, Gottlieb J, Hart E, Middendorf M, Raidl GR (eds) (2002) Applications of evolutionary computing. Springer, Berlin, pp 163–172

Solomon MM (1987) Algorithms for the vehicle routing and scheduling problem with time window constraints. Oper Res 35(2):254–265

Tang H, Miller-Hooks E (2007) Solving a generalized traveling salesperson problem with stochastic customers. Comput Oper Research 34:1963–1987

Toth P, Vigo D (eds) (2002) The vehicle routing problem. SIAM, Philadelphia

Williams TM (1984) Special products and uncertainty in production/inventory systems. Eur J Oper Res 15(1):46–54

Woolridge MJ (2002) An introduction to multiagent systems. Wiley, Chichester

Zäpfel G (1982) Produktionswirtschaft: Operatives Produktions-Management. de Gruyter, Berlin

Zeimpekis V, Giaglis GM (2005) A dynamic real-time fleet management system for incident handling in city logistics. In: Proceedings of 61st IEEE Vehicular Technology Conference (VTC2005 Spring)

Zeimpekis V, Tarantilis CD, Giaglis GM, Minis I (eds) (2007) Dynamic fleet management. Springer, New York

Zeimpekis V, Minis I, Mamassis K, Giaglis GM (2007) Dynamic management of a delayed delivery vehicle in a City Logistics environment. In: Zeimpekis V, Tarantilis CD, Giaglis GM, Minis I (eds) Dynamic fleet management. Springer, New York, pp 197–217