Schedule Nervousness Reduction in Transport Re-Planning

Jörn Schönberger and Herbert Kopfer

Chair of Logistics, Faculty of Business Studies and Economics, University of Bremen
Wilhelm-Herbst-Straße 5, 28359 Bremen, Germany
Email: {jsb; kopfer}@uni-bremen.de

Abstract: This article reports about investigations on a transport scheduling scenario. The generated schedule requires a repeated revision in order to incorporate additionally released requests. Neither the release time nor the location of the associated customer site can be forecasted. Special attention is paid to keep the generated schedules stable, which means that once made scheduling decisions should be maintained (announced arrival times at customer sites, etc.) Initially, several measures for the nervousness degree are proposed. In computational simulations it turns out that a very high nervousness degree appears if the same schedule update strategies are applied for different workloads. As a remedy, the adaptation of the used schedule update strategy to the current schedule performance is proposed. A prototypic algorithm framework is presented and assessed is comprehensive numerical simulation experiments. The adaptation of the schedule update strategies contributes to a significant reduction of some observed nervousness degrees. It reinforces the preservation of once fixed arrival times at customer sites.

Keywords: Adaptive Decision Model, Transport Scheduling, Online Optimization, Uncertainty, Reactive Planning

1. Introduction

A logistic system can be considered as a black-box-system S that transforms a given input signal (requests) into an output signal (logistic processes). A major challenge in the management of logistic systems is to keep the quality of the output on a high and balanced level even if the input signal oscillates with large amplitude. If the system is able to fulfill this property then it is called responsive or systemflexible. Responsiveness and systemflexibility refer to the degree to which changes in the system's environment can be compensated by modifications of the scheduled operations.

In order to cope with the uncertainty of future planning data, the system coordination unit sets up tentative schedules considering all data known at the schedule generation time. The execution of a tentative schedule is started but interrupted immediately if additional planning data are released. Now, previously made operation scheduling decisions are revised in order to integrate the operations associated with the additional requests into the processes. If the frequency of scheduling decision revisions increases then the acceptance of and the trust in such a decision decreases. This planning instability phenomenon is referred to as *schedule nervousness* or *schedule instability* [18].

Interdependence between responsiveness and nervousness is obvious. In order to be responsive it is necessary to revise a given schedule. However, the executed revisions must be chosen very carefully in order to keep the schedule nervousness on a low level. Within this article, a transport system's behavior in a volatile environment is investigated. The management of a demand peak is addressed. Since the transport capacity is limited it is necessary to subcontract requests. A subcontracted request is fulfilled by a hired logistic service provider. The subcontraction acts as a bypass to limit the workload of the own fleet during a demand peak. However, the subcontraction costs of a request are often higher than the expenses for the self-fulfillment of the request. Consequently, this fulfillment mode is not exploited extensively. Since the hired logistic service provider arrive on schedule for each request, it is useful to intensify the subcontractor utilization if the overall punctuality of the request fulfillment decreases or is endangered to decrease.

This article is dedicated to the verification of the following research hypothesis: *The nervousness observed in the transport system is reduced if the utilization degree of the subcontraction request fulfillment mode is adapted to the intermediately detected responsiveness of a transport system.*

The organization of this article is as follows. Section 2 introduces into nervousness classifications, measuring and the relation to flexibility. In Section 3 a re-scheduling scenario from transportation logistics is described which is investigated in order to contribute to the verification of the research hypothesis given above. The applied planning system is outlined in Section 4. Simulation experiments and their results are reported in Section 5.

2. Nervousness and Flexibility

2.1 Previous Work

Decision problems requiring the revision of previously made decisions are introduced as dynamic decision problems [3]. Two generic approaches for coping with the uncertainty of the future are proposed. A-priori-planning exploits probability distributions, which extrapolate the missing data [15]. Reactive-strategies ignore all future data until they becomes definitively known [20].

General ideas of responsiveness and flexibility of value creating systems and networks are introduced in [1], [4], [23] and [22]. Flexibility issues of production systems are surveyed in [29], [10] and [6].

Research on flexibility in the context of transport planning is reported by [7] and [21]. Design and configuration issues for transport networks are discussed but the flexibility aspects of deployment are not addressed.

Planning stability and instability issues are investigated from a general perspective in [17]. Nervousness in inventory management is addressed in [5] and [13] but nervousness in production planning is investigated by Inderfurth and Jensen [14].

2.2 Classification of Nervousness

Nervousness is a symptom appearing during the transition from the so far followed schedule to an updated schedule after additional requests appeared. The first mentioned schedule will be called *preschedule* [16] and the latter one is referred to as *new schedule* in the remainder of this article.

The comparison between a preschedule and the associated new schedule reveal several differences. In the following, the comparison of these two concatenated schedules is discussed.

2.2.1 External and Internal Nervousness

Some revisions of scheduled operations directly affect the customers served by the logistic system S. Typical examples are brought forward or postponed arrival times of transport vehicles at customer sites to pickup or deliver goods or sending repair, maintenance or emergency response teams. Nervousness or instability of data already given to the customers is referred to as *external nervousness*.

Other revisions do not affect the customer or do not receive the customer's attention. The instability of these decisions is called *internal nervousness*. A typical example is the re-assignment of operations to another resource without modification of the operations' completion time.

The prevention of external nervousness is as important as the prevention of internal nervousness. In the first case, the satisfaction of the customers is endangered but in the second case, setup costs and start-up costs occur.

2.2.2 Schedule Transition Nervousness

Let P be a preschedule and Q the associated new schedule. The preschedule consists of a set of operations for which several decisions have been determined at time t_p. For each operation belonging to P, a starting time has been fixed, resource capacities has been reserved (corresponding to the volume that is handled by this operation) and the location(s) where the operation is executed is (are) determined.

All operations which have not been completed before the new schedule Q is set up at time t_Q are also contained in Q. However, the assigned starting time, the volumes of the reserved capacity and the assigned locations have been checked for compatibility with the new requests. If the so far made decisions and the new requests are not compatible then these decisions are revised: an earlier or later starting time has been assigned to one or more operations, the volume to be handled by this operation is increased or reduced and the involved locations are subject of revision.

In case that no decision contained in the preschedule is revised in the new schedule, the preschedule is completely stable. If all decisions made for the preschedule are revised in the new schedule then the preschedule in instable. However, in most of the transitions from a preschedule to a new schedule only a fraction of the operations contained in both consecutive schedules are revised. This observation leads to the following general nervousness definition. Let $N_{P,Q,\mathcal{V}}$ be the number of operations contained in both schedules and let $n_{P,Q,\mathcal{V}}$ be the number of operations subject of a revision during the transition from P to Q using the update preference $\mathcal{V}$. Now, the **schedule nervousness degree $\deg_{\mathcal{V}}(\mathbf{P},\mathbf{Q})$ using update preferences** $\mathcal{V}$ for the transition from P to Q is generally defined as

$$deg_{\mathcal{V}}(P,Q) := \frac{n_{P,Q,\mathcal{V}}}{N_{P,Q,\mathcal{V}}}. \tag{1}$$

Often, only particular decisions are observed during the transition from a preschedule to a new schedule but the definition of the specific schedule nervousness degree remains the same. In transport scheduling the first decision task to be carried out for each schedule revision is the selection of the right fulfillment mode for each request (self-fulfillment and subcontraction). **M**ode **S**election **N**ervousness (MSN) quantifies the variations in the mode decisions between the preschedule P and the new schedule Q. Let $N^{MSN}_{P,Q,\mathcal{V}}$ be the number of all requests contained in P as well as in Q whose fulfillment mode is allowed to be altered. Furthermore, let $n^{MSN}_{P,Q,\mathcal{V}}$ be the number of requests for which the selected fulfillment mode is different in P and Q. Then, the degree of MSN observed in the transition from P to Q applying the update preferences $\mathcal{V}$ is defined as

$$MSN_{\mathcal{V}}(P,Q) := \frac{n^{MSN}_{P,Q,\mathcal{V}}}{N^{MSN}_{P,Q,\mathcal{V}}}. \tag{2}$$

The second decision task in transport schedule generation is the assignment of the operations associated with a request to available resources. Those requests, for which the self-fulfillment has been selected are distributed among the vehicles and a logistic service provider is selected for all subcontracted requests. In operational freight transport there are two kinds of resources: subcontractors and owned vehicles [24]. The decision to subcontract a request cannot be revised but a request that has been assigned to a certain owned vehicle is allowed to be re-assigned to another owned vehicle during a plan update. Let $N^{RAN}_{P,Q,\mathcal{V}}$ be the number of all requests which can be re-assigned from an owned vehicle to another owned vehicles during the transition from P to Q. The expression $n^{RAN}_{P,Q,\mathcal{V}}$ contains the number of requests which have been assigned to different owned vehicles in P and Q and the degree of **R**esource **A**ssignment **N**ervousness (RAN) in the transition from P to Q associated to the update preferences $\mathcal{V}$ is calculated by

$$RAN_{\mathcal{V}}(P,Q) := \frac{n^{RAN}_{P,Q,\mathcal{V}}}{N^{RAN}_{P,Q,\mathcal{V}}}. \tag{3}$$

Both measures MSN as well as RAN quantify internal nervousness. Neither a mode change nor a resource variation might be of interest for the associated customers. In both cases it is possible to keep a once announced request completion time, so that the synchronization of the transport processes with the internal processes of the customers is preserved. In contrast, the variation of announced request completion times represents an external nervousness issue. The symbol $N^{ATN}_{P,Q,\mathcal{V}}$ denotes the number of all requests for which an arrival time shifting is allowed. If $n^{ATN}_{P,Q,\mathcal{V}}$ represents the number of the requests whose completion time is revised during the transition from P to Q, then the **A**rrival **T**ime **N**ervousness (ATN) using $\mathcal{V}$ is defined by

$$ATN_{\mathcal{V}}(P,Q) := \frac{n^{ATN}_{P,Q,\mathcal{V}}}{N^{ATN}_{P,Q,\mathcal{V}}}. \quad (4)$$

2.2.3 Transport System Nervousness

MSN, RAN and ATN describe and quantify the instability of a particular decision with respect to a specific schedule update. A more general quantification is necessary to describe the stability / instability of the system S during a longer period T in which several updates are carried out. Instead of observing and counting the number of revised decisions during the transition from P to Q (at a given time t_Q), it is necessary to consolidate the executed schedule revisions during the generation of the concatenated sequence of schedules $P_i, P_{i+1}, ..., P_{i+k}$ whose update times fall into T. Let $M_{S,T,\mathcal{V}}$ denote the number of all update decisions that must be made during period T and let $m_{S,T,\mathcal{V}}$ be the number of all changes of a decision during the transition from a preschedule to a new schedule during the period T. Then, the **degree $\mathbf{deg^{sys}_{\mathcal{V}}(S,T)}$ of nervousness of system S during period T** using update preferences V is calculated by

$$deg^{sys}_{\mathcal{V}}(S,T) := \frac{m_{S,T,\mathcal{V}}}{M_{S,T,\mathcal{V}}}. \quad (5)$$

The system nervousness degree expresses the inability of the system to maintain and preserve once made decisions during subsequent schedule revisions.

$MSN^{sys}_{\mathcal{V}}(S,T)$, $RAN^{sys}_{\mathcal{V}}(S,T)$ and $ATN^{sys}_{\mathcal{V}}(S,T)$ denote the degree of system nervousness with respect to the mode selection, the resource assignment and the operation sequencing (scheduling). They are defined as described generally in Eq. (5).

2.3 Flexibility and Nervousness of Logistic Operations

Flexibility addresses the ability to integrate additional input into the systems configuration, so that the requirements of the additional input are met (responsiveness). It is distinguished between planflexibility and systemflexibility [27].

Planflexibility quantifies the responsiveness of the system S with respect to update a given preschedule (preplan) P at time t using the integration preferences $\mathcal{V}$ so that the requirements of additionally released requests are met by the new schedule Q. The portion of all requests contained in Q for which the associated requirements are met is called the planflexibility degree. Only those N_t requests, whose completion time falls into the period $[t-\Delta t; t+\Delta t]$, are checked for compliance with the request requirements in order to determined the number n_t of requests completed in the aforementioned period, whose requirements are met. In order to avoid an overweighting of a particular *difficult to integrate request*, all recently completed requests are checked for their compliance with the associated requirements. On the other hand, requests whose scheduled completion time according to the new schedule Q is far in the future are fixed only tentatively. Their completion times are expected to be revised in subsequent schedule updates.

Now the planflexibility degree $F_{\mathcal{V}}(P,t,E)$ of schedule P with respect to the integration of the set of E additional requests using the integration rules $\mathcal{V}$ is defined as

$$F_{\mathcal{V}}(P,t,E) := \frac{n_t}{N_t}. \quad (6)$$

Planflexibility expresses the ability to adapt a particular plan to specific circumstances, but systemflexibility refers to the relative frequency that the considered system is able to integrate additional input E released during a period T so that the requirements of the input are met if the integration preferences $\mathcal{V}$ are used.

Let M denote the number of all possible requests belonging to E and appearing during the period T. If system S can handle m requests appearing during T as required applying the decision preferences V to integrate the requests then the systemflexibility degree of S during T is defined by

$$F^{sys}_{\mathcal{V}}(E,T) := \frac{m}{M}. \quad (7)$$

This degree expresses the *output flexibility* [9] and enables a systemflexibility quantification without having analyzed the internal system structure.

At first glance, a higher systemflexibility degree suggests a higher system nervousness degree since the higher responsiveness requires intensified schedule revisions. However, there are indicators suggesting a contrary interdependence if more requests are forwarded to some logistic service providers. Then, there is a reduced need for updating the routes of the own vehicles so that at the end a less number of scheduling decisions requires a revision. To clarify this issue, the same transport system as used in [27] is investigated in this article with special attention paid now to the dependencies between the achieved systemflexibility degree and the observed system nervousness degree. With the vocabulary introduced in this section, the initial research hypothesis can be refined: *If the systemflexibility degree increases then the nervousness degrees $MSN^{sys}_{\mathcal{V}}(S,T)$, $RAN^{sys}_{\mathcal{V}}(S,T)$ and $ATN^{sys}_{\mathcal{V}}(S,T)$ decrease, if the utilization frequency of subcontraction is adapted to the intermediately observed planflexibility degree instead of using a cost-oriented subcontraction utilization.*

3. Investigated Scenario

We investigate a dynamic decision problem from freight transportation and service providing logistics. Transport resources (vehicles) operating in a restricted area are waiting for their assignment to consecutively released customer requests. Such a request represents the demand to visit a customer location and the simultaneous satisfaction of a time window side requirement. A correct forecast of the release times as well as of the associated customer locations is impossible.

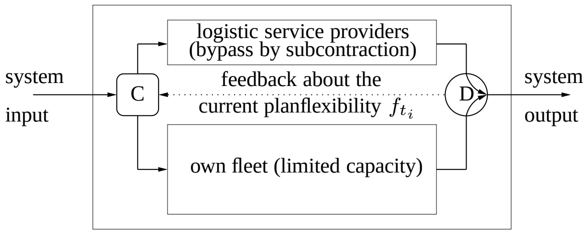

Fig. 1. Components, material flow (solid arcs) and information flow (dotted arcs) in the freight transport scenario.

3.1 Previous and Related Work

A recent survey of dynamic transport scheduling problems is provided by the book of Zeimpekis et al. [31]. A survey of potentials, models and algorithms for the support of subcontraction in freight transport is provided by Krajewska [19]. General layouts of transport disposition systems are discussed in [8]. Bierwirth [2] propose the adaptation of a decision model (*image modification*) in order to improve the systems' reactivity. In the context of the manipulation of a freight transport process optimization model both the variation of the objective function [25], [12], [11] as well as the adjustment of the constraint set have been assessed [26].

3.2 Verbal Description

Road transport systems are important service providing networks for today's value creating systems. They bridge spatial differences between offer and demand and provide spatially scattered distributions or collections of goods. Since nearly 20 years, the demand for road transport has been increased significantly each year as a result of the liberalization and integration of the European markets. However, more and more producing companies have outsourced their transport departments, which act now for own responsibility on a very competitive market. Customers claim least cost transport services and the forwarding companies have reduced their own capacities in the last years so that they are not prepared anymore to serve spontaneous demand peaks with own equipment. Instead they hire other forwarding companies and subcontract transport volume in order to bypass bottleneck situations caused by workload peaks.

Fig. 1 shows the components of the transport system. The own fleet of limited capacity is given by the lower rectangle and the subcontractors (the bypass) are represented by the upper rectangle. The system controller (C) receives the input (additional requests) and decides about the fulfillment mode selection. The mode selection decision is made by comparing the costs for both modes and the least cost mode is selected. An output evaluation component (D) evaluates the current request fulfillment performance in particular it calculates the current planflexibility degree using information about the current reliability of the transport processes (system output). The planflexibility degree f_{t_i} is fed back to the controller (C) who intensifies or thin out the utilization of the subcontractors. The solid arcs represent the material flow in the system and the dotted arc the information flow.

3.3 Bypass Control

The selection of the fulfillment mode for each request (self-fulfillment or subcontraction) is made simultaneously with the route generation decisions. All necessary decisions are made so, that the overall costs for the integration of the additional requests are as small as possible. If the system controller (C) does not intensify or thin out the utilization of subcontraction then no proactive bypass control takes places **(NONE)**.

In previous research reports, two rules for adjusting the utilization degree of subcontraction to the intermediately detected planflexibility degree have been proposed and successfully assessed. Both approaches manipulate the maintained formal mathematical optimization model used for the deployment decisions. A sufficiently high percentage of the incomplete requests is then subcontracted in the next schedule revision.

Adaptation of the Constraint Set (CSAD=Constraint Set ADaptation). A constraint which preselects the subcontraction mode is sharpened (affects more requests) if the planflexibility degree f_{t_i} decreases or if it runs into danger to fall below the threshold value p^{target}. A larger number of requests are put into the set $R(t_i, f_{t_i})$ containing all requests with a predetermined subcontraction decision. In case that the planflexibility degree re-increases the number of requests put in $R(t_i, f_{t_i})$ is reduced [25].

Adaptation of the Objective Function (SDAD=Search Direction ADaptation). The cost coefficients ω_{t_i} and μ_{t_i} weighting both fulfillment modes in the objective function F_{t_i} are re-calculated before the next instance of the decision model is solved: If the planflexibility degree f_{t_i} is low then the cost coefficient ω_{t_i} of the self-fulfillment mode is increase relatively to the cost coefficient μ_{t_i} of the subcontraction mode. If a re-increase of the planflexibility degree is observed then the self-fulfillment cost coefficient ω_{t_i} is lowered step-by-step until it reaches the same value then μ_{t_i} [28].

3.4 Deployment

The transport system hosts 25 own vehicles. Customer demands are released every 100 time units. The average number of incoming requests is 50 but a temporal demand peak of again 100 additional requests during time 1500 until 1700 is simulated.

The system control unit is able to hire a subcontractor for each request. A subcontraction decision cannot be revised. The time window is associated with each request should be considered. If a vehicle or a subcontractor arrives before the time window opens, it has to wait. An arrival after the closure of the time window causes a penalty payment. A subcontractor never arrives late.

The disposition task of the transport system control unit is modeled as an online optimization model with the instances $P_0, P_1, \ldots$. The instances are solved consecutively at the dispatching times $t_0, t_1, \ldots$ Each instance P_i is compiled from all data known at time t_i. A once generated solution TP_i (set of processes) is executed until additional requests arrive at time t_{i+1}. At this time, the transport system control unit interrupts the execution of the so far followed processes. A new optimization model instance $M(t_{i+1})$ is set up applying CSAD or SDAD. After the data of the recent decision task has been collected at time t_i, a mathematical decision model (optimization model) $M(t_i)$ is stated as shown in Eq. (8)-Eq. (13). The requests in the set $R(t_i)$ are known at time t_i and must be distributed among the own fleet W and the incorporated subcontractor(s). The solving of the model Eq. (8)-Eq. (13) completes the schedule update. The new schedule TP_{i+1} replaces the so far not executed process parts from the preschedule TP_i.

All paths p assignable to a vehicle v from the own fleet W are collected in the set $P(t_i)$, all paths executable by a given vehicle $v \in W$ are collected in the set $P_v(ti)$, such a path starts at the current position of vehicle v and ends in the central depot. The binary parameter a_{rp} is 1 if and only if path p serves request r. All requests, for which the SC-mode has already been selected in the preschedule TP_{i-1} are collected in $R^E(t_i)$. The set $R^S(t_i)$ contains all those requests whose on-site operations have already been started but not yet finished at t_i. If r has already been contained in the preschedule and if this request has not yet been subcontracted then $v(r)$ refers to the vehicle which was selected to visit the associated customer site according to the preschedule TP_{i-1}.

The binary decision variable y_r is 1 if and only if request r is subcontracted. Furthermore, we have to assign a (possible empty) path p to a vehicle v and p is assigned to v if and only if the binary decision variable x_{pv} is 1. To evaluate the decisions made, we calculate the travel costs $C^1(p)$ of the path p assigned to an own vehicle of the fleet W, the penalty payments $C^2(p)$ for late arrivals associated with path p and the subcontraction costs $C^3(r)$ of request r.

$$F_{t_i} : \omega_{t_i} \sum_{p \in P(t_i)} \sum_{v \in W} \left(C^1(p) + C^2(p)\right) x_{pv} + \mu_{t_i} \sum_{r \in R(t_i)} C^3(r) y_r \rightarrow \min \qquad (8)$$

$$\sum_{p \in P_v(t_i)} x_{pv} = 1 \; \forall v \in W \qquad (9)$$

$$x_{pv} = 1 \; \forall p \notin P_v(t_i), v \in W \qquad (10)$$

$$y_r + \sum_{p \in P(t_i)} \sum_{v \in W} a_{rp} x_{pv} = 1 \; \forall r \in R(t_i) \qquad (11)$$

$$y_r = 1 \; \forall r \in R^E(t_i) \cup R(t_i, f_{t_i}) \qquad (12)$$

$$\sum_{p \in P_{v(r)}(t_i)} a_{rp} x_{pv(r)} = 1 \; \forall r \in R^S(t_i) \qquad (13)$$

All decisions are made so, that the sum of costs calculated by using the currently applied objective function F_{t_i} is minimized Eq. (8). Each vehicle serves exactly one path Eq. (9) and it is not allowed to assign a path p to a vehicle that cannot serve p Eq. (10). Every request is either subcontracted or served by a vehicle from fleet W Eq. (11). However, previously subcontracted requests remain subcontracted Eq. (12). A request whose on-site fulfillment has already been started but not finished at time t_i cannot be re-assigned to another vehicle Eq. (13).

The objective function Eq. (8) and the constraint Eq. (12) are able to be adapted to the current planflexibility degree f_{t_i}. Thereby, the knowledge acquired during the online-model processing is automatically fed back into the formulation of the model representing the next decision task.

4. Algorithmic Approach

The algorithmic framework deployed to cope with the online-optimization model introduced in Section 3 is shown in Fig. 2. It was originally presented in [27].

Initially, the iteration counter i is set to 0 (a) and the first planning time t_i is fetched (b). Next, an initial solution is generated (c) and broadcasted to the vehicles and to the subcontractors (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 requests become known. 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 collected in the set $R^+(t_i)$ (i). Next, it is checked whether the consideration of the additional requests compromises the current processes (j). The procedure falls back into an idle state if no process corruption occurs. Otherwise, the current planflexibility f_{t_i} is calculated (k) and the intervention intensity s_{t_i} (describing the severeness of the manipulation) is determined (l). If CSAD is applied then the set $R(t_i, f_{t_i})$ is formed which contains all those requests, which are prematurely directed into the subcontraction fulfillment mode. Otherwise if SDAD is used then the coefficients $(\omega_{t_i}, \mu_{t_i})$ defining the next applied objection function F_{t_i} are determined (m). Afterwards, the new decision model $M(t_i)$ is defined (n) and a high quality solution of this model is derived (o) to replace the so far followed solution. For the derivation of the new solution we start the Memetic Algorithm developed in [24]. The new solution is broadcasted to inform the field teams, the subcontractors and the customers (p). Again, the procedure falls back into the idle (waiting) state (q).

The procedure GET_INTERVENTION_INTENSITY(f_{t_i}) evaluates the current planflexibility degree and returns the control value s_{t_i}, which is normalized between 0 and 1. If s_{t_i} is zero then no model adaptations are applied but if s_{t_i} equals 1 then all possible adjustments are implemented. The control value s_{t_i} is 0 if the current planflexibility degree f_{t_i} is higher than 85% but it is 1 if the current planflexibility degree f_{t_i} has fallen below 75%. If the planflexibility degree increases (decreases) between 0.75 and 0.85 then s_{t_i} is proportionally lowered (enlarged).

The control value is used to setup the necessary decision model adaptations. If CSAD is used to adapt the decision model Eq. (8) - Eq. (13) then the ordered pair of the coefficients of the objective function is fixed to (1,1) but the set $R(t_i, f_{t_i})$ must be compiled for every re-planning task. At first $R(t_i, f_{t_i})$ is emptied. Secondly, the number $n(t_i)$ of requests to be inserted into $R(t_i, f_{t_i})$ is set to $\lceil s_{t_i} \cdot \mid R^+(t_i) \mid \rceil$. Finally, $n(t_i)$ requests are randomly selected from $R^+(t_i)$ and inserted into $R(t_i, f_{t_i})$.

In case that SDAD is incorporated for the decision model adjustment the two coefficients ω_{t_i} and μ_{t_i} are determined previously to the solving of the model $M(t_i)$. The first coefficient ω_{t_i} is fixed to the value 1 but μ_{t_i} is adapted to the current value

```
PROCEDURE process_management();
(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)   f_{t_i} := GET_CURRENT_PLANFLEXIBILITY(t_i);
(l)   s_{t_i} := GET_INTERVENTION_INTENSITY(f_{t_i});
(m)   R(t_i, f_{t_i}) := SPECIFY_INTERVENTION(s_{t_i},R+(t_i)); (only CSAD)
      (ω_{t_i}, μ_{t_i}) := SPECIFY_COEFFICIENTS(s_{t_i}); (only SDAD)
(n)   M(t_i) := DEFINE_MODEL(t_i, CurrentSolution,R+(t_i),R(t_i, f_{t_i})); (only CSAD)
      M(t_i) := DEFINE_MODEL(t_i, CurrentSolution,R+(t_i),(ω_{t_i}, μ_{t_i})); (only SDAD)
(o)   CurrentSolution := SOLVE_MODEL(M(t_i));
(p)   BROADCAST(CurrentSolution);
(q)   goto (e);
(r)   stop();
```

Fig. 2. Pseudo code of the algorithmic framework

of the control signal s_{t_i}. Let K denote the average quotient between the fulfillment costs in the subcontraction mode and the fulfillment costs in the self-fulfillment mode. The coefficient μ_{t_i} is set to $1 + s_{t_i} \cdot K$. If the current planflexibility degree is quite high then the control signal s_{t_i} is 0 and $\mu_{t_i} = 1$. If the current planflexibility degree is low then the control signal is close to 1 and $\mu_{t_i} \approx 1 + K$, so that the objective function (Eq. 8) recognizes the subcontraction as the cheaper fulfillment mode.

5. Numerical Experiments

5.1 Experimental Setup

For the assessment of the previously introduced bypass control strategies comprehensive numerical experiments have been setup and carried out. Artificial test instances introduced in [25] are used. Streams of consecutively arriving requests are taken from the Solomon instances [30] $P \in \{R103, R104, R107, R108\}$. In these streams, 50 additional requests are released every 100 time units. During the period from t=1500 until t=1700, an additional workload of 100 requests is released leading to the overall number of 150 additional requests during this period.

The three bypass control strategies $\mathcal{V} \in \{NONE, SDAD, CSAD\}$ have been applied to the four request sets. Every simulation run has been seeded with three different values. Overall, numerical results from $3 \cdot 4 \cdot 3 = 36$ simulation runs are reported in this section.

After a startup phase of 1000 time units, several performance indicators have been recorded for the next 4000 time units. A schedule revision is performed every 100 time unit as response to the arrival of the additionally released requests.

In order to enable a thorough analysis of the simulation results, the following performance indicators have been recorded during the experiments: The averagely observed values for $MSN_{\mathcal{V}}$, $RAN_{\mathcal{V}}$ and $ATN_{\mathcal{V}}$ have been calculated for each bypass adaptation strategy $\mathcal{V}$ as well as the average of the planflexibility degree $F_{\mathcal{V}}$. After a simulation experiment has been completed the achieved degrees for the system mode selection nervousness ($MSN_{\mathcal{V}}^{sys}$), the system resource assignment nervousness ($RAN_{\mathcal{V}}^{sys}$) and the system arrival time nervousness ($ATN_{\mathcal{V}}^{sys}$) as well as for the systemflexibility degree $F_{\mathcal{V}}^{sys}$ have been calculated. In the next subsection, the averagely observed values are reported.

5.2 Results

Fig.3 (continuous curve) shows that the planflexibility degree collapses significantly after the demand peak occurs at time 1500 until time 1700 (NONE). Both adaptive strategies SDAD (dashed curve) as well as CSAD (dotted curve) contribute to the stabilization of the planflexibility degree so that the planflexibility degree decreases only slightly and for a short period after the demand peak occurrence if SDAD or CSAD are incorporated.

The second column in Table 1 summarizes the averagely observed systemflexibility degrees broken down into the three integration strategies. In the reference experiment without integration preference adaptation (NONE) a systemflexibility degree of 74.1% is reached. If the constraint set is adapted to the current planflexibility degree (CSAD) then the significantly increased systemflexibility degree 83.3% is achieved. An even slightly better systemflexibility degree of 83.8% is observed if the objective function is parameterized adaptively (SDAD).

In order to compare the systemflexibility and the degree of nervousness, the averagely observed degrees for mode selection, resource assignment and arrival time are presented in the columns 3-5 in Table 1. Since both adaptive strategies CSAD and SDAD enforce and intensify the utilization of the subcon-

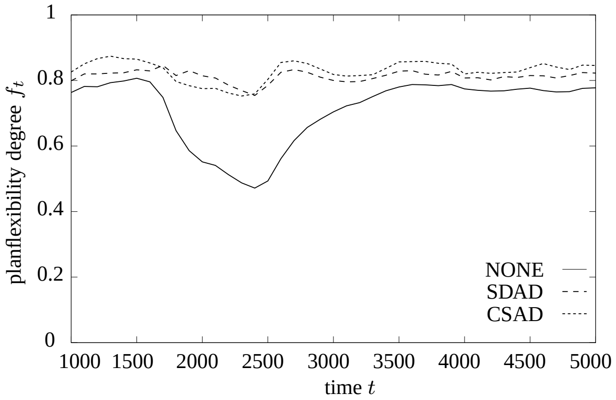

Fig. 3. Evolution of the planflexibility degree f_{t_i}

Table 1. Observed system nervousness values

$\mathcal{V}$	$F_{\mathcal{V}}^{sys}$	$MSN_{\mathcal{V}}^{sys}$	$RAN_{\mathcal{V}}^{sys}$	$ATN_{\mathcal{V}}^{sys}$
NONE	74.1%	1.35%	35.5%	56.8%
CSAD	83.3%	5.25%	28.5%	38.2%
SDAD	83.8%	6.25%	31.4%	48.8%

Table 2. Frequency of requests which have been proponed or deferred compared to their initial completion time

$\mathcal{V}$	earlier	later	unvaried
NONE	16%	41%	43%
SDAD	17%	32%	51%
CSAD	10%	28%	62%

traction mode, an increase of $MSN_{\mathcal{V}}^{sys}$ from 1.35% (NONE) to 5.25% (CSAD) respectively 6.25% (SDAD) is observed. Thus, the research hypothesis cannot be verified for this particular internal system nervousness degree.

A different observation is made for the resource assignment system nervousness degree: RAN_{NONE}^{syst} is 35.5% but this degree decreases down to 31.4% (SDAD) and even 28.5% (CSAD) if the integration preferences are adapted to the intermediately observed planflexibility degree. The conclusion of this observation is that the research hypothesis is verified for this specific nervousness degree.

With respect to the degree of the external arrival time nervousness, the observed results enable a clear verification of the research hypothesis. If the knowledge of the intermediate planflexibility degree is not exploited for the variation of the integration preferences (NONE) then the initially announced request completion time is revised for 56.8% of all incoming requests. If SDAD is applied then the arrival time revision degree is reduced below 50% to ATN_{SDAD}^{sys}=48.8%. In case that CSAD is used to adjust the next decision model a further reduction below 40% is achieved: $ATN_{\mathcal{V}}^{sys}$=38.2%. It is concluded, that the research hypothesis is verified for the arrival time nervousness.

The decrease of the ATN^{sys}-values is significant, so that this particular aspect has been analyzed in more detail. The second column in Table 2 contains the percentage of requests whose final completion time is earlier than their initially announced completion time. A reduction of the percentage of

Table 3. Evolution of the systems workload

$\mathcal{V}$	$WL_{\mathcal{V}}^{max}$	$PD_{\mathcal{V}}$
NONE	+287%	39%
SDAD	+208%	20%
CSAD	+232%	23%

this *left shifting* is observed only if CSAD is applied. From the values presented in the third column, it is concluded that both adaptive strategies support the prevention of deferments. The postponement percentage for a request decreases from 41% (NONE) to 31% (SDAD) respectively 28% (CSAD). Altogether, the percentage of un-rescheduled requests increases from 43% (NONE) to 51% for the SDAD-application and even up to 62% if CSAD is used.

The system's workload (expressed in terms of waiting requests) is finally analyzed. Independently from the applied bypass control strategy the number of waiting requests increases significantly after the demand peak's introduction. However, the maximal number of waiting requests ($WL_{\mathcal{V}}^{max}$) is influenced by the applied bypass controll strategy (Table 3): If NONE is applied then the number of waiting (known but not yet completed requests) increases maximally up to 287%, if CSAD is applied then a lifting of 232% takes place but if SDAD is incorporated than an increase of 208% is observed. The percentage of the observation time period in twhich a workload is detected that is higher than the average pre-peak workload is stored in $PD_{\mathcal{V}}$. If the punctuality feedback information is not exploited (NONE) then the pre-peak workload is exceeded after the peak's appearance during 39% of the observation intervall [1000;5000]. The application of CSAD reduces $PD_{\mathcal{V}}$ down to 23% and the deployment of SDAD let $PD_{\mathcal{V}}$ fall down to 20%.

The observed numerical results do not verify the research hypothesis in general. However, it has been observed that the adaptation of the integration preferences to the intermediate planflexibility degree supports the reduction of the external nervousness degree of the arrival times at customer sites. The bypass control by feedback information about the process quality is a reasonable tool to achieve a higher reliability with respect to the announced arrival times so that the customers' trust in a once submitted arrival time increases.

6. Conclusions

An adaptive control of a workload bypass to balance the workload entering a fleet of vehicles of limited capacity has been proposed and investigated. The control has been realized by adjusting the utilization of the expensive subcontraction (the bypass) to the intermediately observed planflexibility degree, which represents the current transport system responsiveness.

It has been demonstrated that the balancing of the workload reduces the system nervousness in several aspects. Therefore it contributes to a stabilization of once made scheduling decisions.

Future research will be dedicated to the transfer of adaptive control strategies to other logistic systems which have to cope with oscillating workload. Furthermore, additional performance criteria will be combined in the control signal determination, so that different aspects of process performance can be considered in the adjustment of the schedule update preferences and in particular to the adjustment of the maintained decision model.

Acknowledgement

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

References

[1] M Barad and D E Sapir, Flexibility in logistic systems – modeling and performance evaluation. International Journal of Production Economics, No. 85, 2003, pp. 155–170.

[2] C Bierwirth, Adaptive Search and the Management of Logistics System, Kluwer, 1999.

[3] B Bremer, Dynamic decision making: Human control of complex systems. Acta Psychologica, Vol. 81, 1992, pp. 211–241.

[4] X de Groote, The flexibility of production processes: A general framework. Management Science, Vol. 40, No. 7, 1994, pp. 1587–1617.

[5] T de Kok and K Inderfurth, Nervousness in inventory management: Comparison of basic control rules. European Journal of Operational Research, No. 103, 1997, pp. 55–82.

[6] A de Toni and S Tonchia, Manufacturing flexibility: a literature review. International Journal of Production Research, Vol. 36, No. 6, 1998, pp. 933–945.

[7] E Feitelson and I Salomon, The implications of differential network flexibility for spatial structures. Transportation Research Part A, Vol. 34, 2000, pp. 449–479.

[8] S Gayalis and I Tatsiopoulos, Design of an it-driven decision support system for vehicle routing and scheduling. European Journal of Operational Research, Vol. 152, No. 2, 2004, pp. 382–398.

[9] R W Grubbström and J Olhager, Productivity and flexibility: Fundamental relations between two major properties and performance measures of the production system. International Journal of Production Economics, Vol. 52, 1997, pp. 73–82.

[10] Y P Gupta and S Goyal, Flexibility of manufacturing systems: Concepts and measurements. European Journal of Operational Research, Vol. 43, 1989, pp. 119–135.

[11] K Gutenschwager, Online-Dispositionsprobleme in der Lagerlogistik, Physica-Verlag, 2002.

[12] K Gutenschwager, F Böse and S Voß, Effiziente Prozesse im kombinierten Verkehr - ein neuer Lösungsansatz zur Disposition von Portalkränen. Logistik Management, Vol. 5, No. 1, 2003, pp. 62–73.

[13] K Inderfurth, Nervousness in inventory control: analytical results. OR Spektrum, Vol. 16, 1994, pp. 113–123.

[14] K Inderfurth and T Jensen, Planungsnervosität im Rahmen der Produktionsplanung. Zeitschrift für Betriebswirtschaft, Vol. 67, No. 8, 1997, pp. 817–843.

[15] P Jaillet, A priori solution of travelling salesman problem in which a random subset of customers are visited. Operations Research, Vol. 36, No. 6, 1988, pp. 929–936.

[16] M T Jensen, Robust and Flexible Scheduling with Evolutionary Computation, PhD thesis, University of Aarhus, Denmark, 2001.

[17] T Jensen, Planungsstabilität in der Materiallogistik, Physica-Verlag, 1996.

[18] S N Kadipasaoglu and S V Sridharan, Measurement of instability in multi-level mrp systems. International Journal of Production Research, Vol. 35, No. 3, 1997, pp. 713–737.

[19] M Krajewska, Potentials for efficiency increase in modern freight forwarding, Gabler, 2008.

[20] S O Krumke, Online-Optimization - Competitive Analysis and Beyond, Habilitation Thesis. Technical University Berlin, Germany, 2001.

[21] E K Morlok and D J Chang, Measuring capacity flexibility of a transport system. Transportation Research Part A, Vol. 38, 2004, pp. 405–420.

[22] R Pibernik, Flexibilitätsplanung in Wertschöpfungsnetzwerken, DUV, 2001.

[23] C Schneeweiß and H Schneider, Measuring and designing flexibility as a generalized service degree. European Journal of Operational Research, No. 112, 1999, pp. 98–116.

[24] J Schönberger, Operational Freight Carrier Planning, Springer, 2005.

[25] J Schönberger and H Kopfer, Flexible transport process planning in volatile environments and the adaptation of a cost-based objective function, In Logistikmanagement, A Otto and R Obermaier, Eds. DUV, 2007, pp. 263–283.

[26] J Schönberger and H Kopfer, On decision model adaptation in online optimization of a transport system, In Management logistischer Netzwerke, H O Günther, D C Mattfeld, and L Suhl, Eds. Physica-Verlag, 2007, pp. 361–381.

[27] J Schönberger and H Kopfer, Flexibility of logistic systems – responsiveness improvement of a transport system, Tech. rep., University of Bremen, 2008.

[28] J Schönberger and H Kopfer, Manipulating the decision behavior of a subordinate service centre by a closed-loop cost accounting scheme control, In Information- und Kommunikationssysteme in Supply Chain Management, Logistik und Transport, L Suhl, D C Mattfeld, H O Günther, and S Voß, Eds. DSOR Contributions to Information Systems, 2008, pp. 3–19.

[29] A K Sethi and S P Sethi, Flexibility in manufacturing: A survey. International Journal of Flexible Manufacturing Systems, Vol. 2, 1990, pp. 289–328.

[30] M Solomon, Algorithms for the vehicle routing problem with time window constraints. Operations Research, Vol. 35, No. 2, 1987, pp. 254–265.

[31] V Zeimpekis, C D Tarantilis, G M Giaglis and I Minis, Eds., Dynamic Fleet Management, Springer, 2007.