COORDINATION PRODUCTION AND TRANSPORTATION SCHEDULING IN THE SUPPLY CHAIN

Kadir Ertogral, S. David Wu and Laura I. Burke
The Manufacturing Logistics Institute
Department of Industrial and Manufacturing Systems Engineering
Lehigh University, Bethlehem, Pennsylvania

ABSTRACT

Coordinating operational and logistic functions across facilities and companies is key to supply chain integration. In this paper, we consider the coordination of production and transportation scheduling, a crucial link of this integration. In manufacturing-centric industries such as automotive and electronics, transportation costs constitute the second largest cost component following the production costs. While the potential of integrating production and transportation planning could be significant, there has been a lack of research and practice in integrating these functions. Functional silos or ad hoc interfaces are commonly seen in the industry. In this research, we propose an analytic study that investigates the effects of integrating production and transportation planning. We study integrated optimization models that reconcile the viewpoints from transportation and production planning and analyze the costs introduced by coordination. Using a Lagrangean decomposition scheme, we demonstrate computationally the nature of the comprise between production and transportation decisions, and the value of integration.
1. Introduction

Supply chain management attracts central attention in the 1990's as firms start to realize the potential cost benefits of integration: across facilities, across layers of the product/functional hierarchy, or across supply tiers in the production process. Integration of a manufacturing supply chain involves myriad planning, control and coordination activities across layers of facilities which transform raw material into final goods. The driving force behind supply chain integration is the ever-increasing market pressure on responsiveness, cost efficiency, and the drastically shorting product life-cycle. This integration is expected to bring significant cost improvement by utilizing resources more efficiently throughout the chain while improving customer service. In this research, we will tackle a critical segment of this integration focusing on the integration of production and transportation planning in manufacturing supply chains typical in automotive and electronic industries.

Main cost factors within a supply chain can be put into the categories of production, transportation, inventory, and material handling costs. The composition of these operational costs relative to total costs varies largely by industry. However, production cost is the largest of all in almost all the industries, followed by transportation and inventory costs (Chen (1997)). Despite of this, historically production and transportation logistics have been dealt with separately both in industry and academia. In industry, a production plan is developed and then a transportation plan is worked out by either the transportation department of a company, or a third party transportation provider, who adheres to an established shipping plan aiming at reducing transportation costs. Transition between the two functions relies on inventory buffers of different forms, e.g., in the automotive industry, a ten- to fourteen-day inventory buffer is a common practice for the very purpose. The extent to which the transportation costs are considered in the production plan does not go beyond a simple evaluation of a few transportation channels. Materials requirements planning, the most common decision support tool for production planning in the industry, does not even consider the capacity restrictions of production resources, let alone the transportation cost.

Parallel to industry practice, researchers in academia have approached the two problems separately. There has been an enormous body of research on production planning models, and in specific, on lot-sizing models. The main tradeoff considered here is between the inventory carrying and setup costs. For the more recent multi-level, multi-item, multi-
period dynamic lot-sizing problem (MLMILP), most research focuses on the development of heuristics due to computational requirements for practical size problems. On the other hand, for transportation planning problems, *vehicle routing and scheduling* has also been studied extensively in the literature: ranging from the basic traveling salesman problem to multi-vehicle pickup and delivery problem with time windows (m-PDPTW). To study the main trade-off between production and transportation decisions we will investigate means of integrating decisions characterized by MLMILP and m-PDPTW type models. We propose new models for both part of the problems that are tailored to integrated decision making and a Lagrangean decomposition approach for the solution of the integrated problem.

The literature on integrated production and transportation study has been rather limited. Chandra and Fisher (1994) showed empirically the value of integrating production and transportation decisions in an environment which involves a single production facility and multiple customers. They report gains ranging between 3% to 20% obtained by integrating production planning and vehicle routing in a heuristic manner. Researchers have proposed various extensions of production or transportation model to incorporate limited aspects of the other sides. We will review, in the following sections the literature for MLMILP and m-PDPTW models.

1.1 Related Literature

**Multi-level multi-item dynamic capacitated lot sizing problem**

MLMILP can be defined as follows: given the external demand for end items over a time horizon of $T$ periods, determine a *production plan* that minimizes total inventory holding and setup costs, subject to capacity restrictions of resources, without end items backlogging. There have been several formulations of the problem in the literature (Stadtler 1996). For the purpose of later discussion, we will provide a normative formulation of the problem which is similar to the one presented in Tempelmeier et al. (1996). Without loss of generality, we exclude the production cost from the formulation since it will not affect the analysis.

**Notation:**

- $\gamma_{kt}$ : binary setup variable for item $k$ in period $t$. $\gamma_{kt} = 1$ if $q_{kt} > 0$ and $\gamma_{kt} = 0$ o.w.
- $a_{ki}$ : number of units of item $k$ required to produce one unit of item $i$. 


\( b_{jt} \) : available capacity of resource \( j \) in period \( t \).
\( d_{kt} \) : external demand for item \( k \) in period \( t \).
\( h_k \) : inventory holding cost for item \( k \).
\( s_k \) : setup cost for item \( k \).
\( J \) : number of resources.
\( K \) : number of items.
\( K_j \) : set of items that use resource \( j \).
\( T \) : number of periods.
\( M \) : a large number.
\( N_k \) : set of items that are immediate successors of item \( k \) in the product structure.
\( q_{kt} \) : lot size for item \( k \) in period \( t \).
\( t_{bk} \) : production time per unit of item \( k \).
\( t_{rk} \) : setup time of item \( k \).
\( y_{kt} \) : inventory of item \( k \) at the end of period \( t \).
\( z(k) \) : deterministic minimal lead time for item \( k \).

The formulation is as follows:

\[
\begin{align*}
\text{Min} & \sum_{t=1}^{T} \sum_{k=1}^{K} (h_k y_{kt} + s_k y_{kt}) \\
\text{subject to} & \\
y_{kt-1} + q_{kt} - \sum_{i \in N_k} a_{ki} q_{it} - y_{kt} = d_{kt} & k = 1, 2, \ldots, K; \ t = 1, 2, \ldots, T \\
\sum_{k \in K_j} (t_{rk} y_{kt} + t_{bk} q_{kt}) & \leq b_{jt} & j = 1, 2, \ldots, J; \ t = 1, 2, \ldots, T \\
q_{kt} - M y_{kt} & \leq 0 & k = 1, 2, \ldots, K; \ t = 1, 2, \ldots, T \\
q_{kt} & \geq 0, \ y_{kt} & \geq 0, \ y_{kt} \in \{0, 1\} & k = 1, 2, \ldots, K; \ t = 1, 2, \ldots, T
\end{align*}
\]

**Solution approaches**

It has been shown that even finding a feasible solution to the problem is NP complete (Maes et al. 1991). Very few optimal solution methods for the problem in the literature (Pocket et al. 1991, Chapman 1985) had to be quite limited in problem size. Therefore, almost all of the research has been concentrated on heuristic approaches.
The heuristic approaches ranges from Lagrangean relaxation to tabu search. Zahorik et al. (1984) describe an optimization based heuristic approach, employing a 3-period network flow formulation of the problem, for a quite restrictive case with no setup cost or time. Billington et al. (1986) introduce a branch and bound heuristic using Lagrangean relaxation. It is assumed that the capacity restriction exists only at one level in the BOM structure. They use a price-directed approach to reflect both the multilevel structure and the capacity limitations. Maes et al. (1991) address the complexity of finding feasible solutions to capacitated MLMILP and present three similar heuristics for the solution. The three heuristics differ in the way they round the binary setup variables, which are obtained from LP relaxation, up to 1 or down to 0. Roll and Karni (1991) present a heuristic approach which consists of the application of eight different subroutines. These subroutines either convert an infeasible solution to a feasible one or improve a given solution. Maes and Van Wassenhove (1991) extend their ABC algorithm (Maes and Van Wassenhove 1986) for capacitated single level lot sizing problems to multilevel problems with serial BOM structure. Kuik et al. (1993) use two meta heuristic techniques, namely simulated annealing and tabu search, in their heuristics. Search neighborhood of these heuristics is defined on the setup variables. They computationally show that their heuristics perform better on the average compared to those in Maes et al. (1991). Tempelmeier and Helber (1994) address four variants of a two-phase heuristic approach for the problem. In the first phase, the problem is decomposed into a set of single level capacitated multi-item lot sizing problems and an ordering for the solution of these single level problems is determined. In the second phase, the single level problems are solved in the order determined in the first phase by using a heuristic due to Dixon and Silver (1981). Tempelmeier and Derstroff (1996) explain a solution method based on the Lagrangean relaxation of the problem. They relax both inventory and capacity constraints to obtain single item uncapacitated lot sizing problems. Summation of the solutions of these relaxed problems are taken as the lower bound and this lower bound is compared to an upper bound. The upper bound is calculated by solving the single item problems from end items down to lowest level items. At the end, possible capacity violations is eliminated through a shifting procedure. In non of the above approaches, transportation cost is considered.

Stadtler (1996) collectively presents five mixed integer programming formulations of the problem proposed in the literature including two network oriented models, namely simple plant location and shortest route models. He compares these models in terms of problem size, integrality gap of LP relaxation and computation time. Two recent survey papers regarding the problem are Drexl and Kimms (1997), Kuik et al. (1994).
Multiple vehicle pickup and delivery problem with time windows

The multiple vehicle pickup and delivery problem with time windows (m-PDPTW) can be stated as construction of optimal routes for a set of vehicles so as to satisfy a specified set of transportation requests. Each request is described by a pickup location, a delivery location, the amount of load to be transported, and pickup and delivery service time windows. Optimal routes here are ones that optimize a certain objective function. We give a mathematical formulation of the problem in section 2.3.

A typical objective for problems concerning transportation of goods is minimization of the total distance traveled by the vehicles or the total travel time. The minimization of the number of vehicles used or the associated total vehicle fixed cost might be considered together with the above objectives. The m-PDPTW that involves the transportation of people is called the dial-a-ride problem in the literature. Because drivers and vehicles are the most expensive parts in a dial-a-ride system, minimizing the number of vehicles to serve all the requests is commonly the objective. In dial-a-ride systems, minimizing client inconvenience is another usual objective.

m-PDPTW in the context of transportation of goods has not received the same extensive attention as some other vehicle scheduling problems. Due to the large solution spaces inherent in routing problems, heuristic solution methods are necessary to solve practical size problems involving complicating side constraints.

Solution approaches

In all of the following approaches load size to be transported and the transportation time allowances are taken as parameters. The research on m-PDPTW has been concentrated on dial-a-ride type systems. We first discuss the heuristic approaches. Jaw et al. (1986) address an insertion based heuristic algorithm for multiple vehicle dial-a-ride problems. The objective is to minimize a weighted combination of customer inconvenience and system cost. The insertion criterion is very simple: Among all feasible insertion points of a customer to the vehicle schedules, select the cheapest one. If there is no feasible insertion point for the customer, introduce an additional vehicle to the system. Bodin and Sexton (1986) present a cluster first and route second type approach again for dial-a-ride type problems with multiple vehicles. In the first phase, clients are assigned to vehicles and a
single vehicle dial-a-ride problem for each vehicle is solved to get an overall initial solution. This initial solution is improved in the second phase by reassigning clients with bad service to other vehicles. Dumas et al. (1989) develop a heuristic approach for the same type problem. Their heuristic uses the traditional cluster first route second approach but moves a part of the clustering problem into the routing problem. The main idea is to create route segments for small groups of requests, called mini-clusters, then stringing them together to find a solution. Mini-clusters are basically the route segments which starts and ends with an empty vehicle. Ioachim et al. (1991) illustrate a two phase heuristic approach which uses the mini-clustering idea. In the first phase, mini-clustering problem is solved. They use mathematical optimization techniques to define a set of mini-clusters. In the second phase, the problem of linking the mini-clusters to construct routes for vehicles is modeled and solved as m-TSPTW. Destrosiers et al. (1991) suggest a parallel insertion heuristic to create mini-clusters. They base their insertion criterion on the concept of neighboring requests which is defined as the set of request satisfying certain temporal or spatial proximity restrictions. Fisher et al. (1995) describe a heuristic approach for a truck-load type problem where vehicles can not carry more than one load at a time and, therefore, they go to the delivery location directly after visiting the pickup point. The approach utilizes a network flow based relaxation to find near feasible solutions. Then these solutions are made feasible using some simple heuristic techniques. The computational study involving some real problems shows that the approach resulted into solutions within 1% of the optimal.

An optimal algorithm for the problem in the context of transportation of goods is proposed by Dumas et al. (1991). They present a set partitioning formulation for m-PDPTW and use a column generation type method to solve the problem. Since the size of the set of admissible routes can get very large it is practically inefficient to apply exhaustive enumeration. Therefore, in order to solve a linear relaxation of the set partitioning formulation, the authors use the column generation idea to introduce new columns corresponding to admissible routes. Columns in the set of admissible routes are generated as needed, as a solution to the constrained shortest path problem which defines the admissible routes. Ruland and Rodin (1997) explore the polyhedral structure of the dial-a-ride problems and develop four classes of valid inequalities. They use these inequalities in a branch-and-cut type procedure to solve the problem.

Savelsbergh and Sol (1995) describe the general characteristics of the problem and present a survey classified into problem types and solution methods.
2. Problem Description and Formulation

2.1 Problem description

In a broad sense, we can state our problem as the integration of production and transportation planning functions in a supply chain environment so as to minimize the operational cost over these two planning functions. There exists a complementary relationship between production and transportation functions in general. Production functions utilize the services provided by transportation and the price of these services is determined based on the quality of the services required by production. In this research, we try to address this primal-dual type conjunction between production and transportation models.

Traditionally production and transportation functions are separated using large buffers. But there is a constant market pressure for firms to reduce the inventory in their distribution channels and to make themselves operate in a just-in-time manner. It has become a must for firms to explore a closer coordination between production and distribution functions. Although the literature that addresses both production and distribution problems is plentiful, very few models try to combine and solve these problems simultaneously. Among the combined models, even fewer involve routing decisions. The lack of research on coordinating these two activities may be that 1) These two problems by themselves are hard problems and the combined problem is not tractable, 2) Different organizational entities are responsible for the planning of these two activities. There has been no obvious need to combine these problems.

We assume a manufacturing supply chain environment with general product structure. We consider a MLMILP model for the production planning problem. Conventional lot sizing models in the literature do not involve transportation time between production locations in the supply chain. Delivery of goods is generally assumed to be instantaneous. In our problem we consider the transportation times as parameters in the model.

For the transportation portion of our problem, we use a model based on m-PDPTW formulation. In our case we need to construct a different m-PDPTW than the typical one. In a typical m-PDPTW, there are fixed loads to be picked up and delivered and there is a fixed travel time allowance for each request, which we shall define as the time period
between the end of the pickup time window and the start of the delivery time window. Since we are trying to integrate two planning activities, we will be seeking the size of the loads and the travel time allowances that are mutually beneficial to both production and transportation. Therefore, size of the loads and the travel time allowances will be taken as variables in the pickup and delivery part of our problem. This inevitably results in a harder pickup and delivery problem than the typical one.

2.2 Production planning model

Notation used in the model is as follows:

**Variables:**
- $Q_{iti}$ : Amount of production of item $i$ in period $t$.
- $I_t$ : Inventory of item $i$ as finished good at the end of period $t$.
- $CI_{ij,t}$ : Inventory of item $i$ as subassembly to be used in the production of item $j$ at the end of period $t$.
- $PS_{iti}$ : Production start time of item $i$ in period $t$ in terms of time buckets.
- $SL_{ij,t}$ : Size of load of item $i$ which is sent in period $t$ for the production of item $j$.
- $V_t$ : Binary setup variable for item $i$ in period $t$.

**Parameters:**
- $hc_i$ : Inventory holding cost of item $i$ per period when item $i$ is regular inventory.
- $sc_i$ : Setup cost for item $i$.
- $deltah_i$ : Percent increase over $hc_i$ in the inventory holding cost of item $i$ when item $i$ becomes subassembly inventory.
- $ini_i$ : Initial inventory of item $i$.
- $inc_{ci,j}$ : Initial subassembly inventory of item $i$ for the production of item $j$.
- $st_i$ : Setup time for item $i$ in time buckets.
- $pt_i$ : Production time for item $i$ in time buckets.
- $trt_{ij}$ : Traveling time between production location $i$ and $j$ in time buckets.
- $cp_i$ : Amount of capacity consumed when one unit of item $i$ is produced.
- $cs_i$ : Amount of capacity consumed when a setup for item $i$ is performed.
- $lc_i$ : Production location where item $i$ is produced.
- $cap_i$ : Capacity of resource $i$ in period $t$.
- $dem_{iti}$ : Demand for item $i$ in period $t$.
- $a_{ij}$ : Number of item $i$ required to produce one unit of item $j$.
- $per$ : Length of a production planning period in time buckets.
- $M$ : A large number.
Sets:

$T$ : Set of planning periods.
$I$ : Set of items.

$N_i$ : Set of indices of the immediate successor items of item $i$ in the product structure.

$RS$ : Set of resources

$RS_i$ : Set of items that are produced at resource $i$.

$L$ : Set of pairs of indices $(i, j)$ representing the links between items in the product structure.

The proposed production planning model is as follows:

\[
\begin{align*}
\text{Min} & \quad \sum_{i \in T} \left( \sum_{i \in I} (hc_i I_{it} + sc_i V_{it}) + \sum_{(i, j) \in L} (1 + \text{delta}_i) hc_i CI_{ijt} \right) \\
\text{subject to :} & \\
I_{i,t+1} - Q_{i,t} - I_{it} - \text{dem}_{it} - \sum_{j \in N_i} SL_{ij1} & = 0 \quad \forall i \in I, \forall t \in T \setminus \{1\} \\
\text{ini}_i + Q_{i1} - I_{i1} - \text{dem}_{i1} - \sum_{j \in N_i} SL_{ij1} & = 0 \quad \forall i \in I \\
CI_{ij,t-1} + SL_{ij,t-1} - CI_{ijt} - a_{ij} Q_{ij} & = 0 \quad \forall (i, j) \in L, \forall t \in T \\
\text{inci}_{ij} - CI_{ij1} - a_{ij} Q_{i1} & = 0 \quad \forall (i, j) \in L \\
(t-1)\text{per} + \text{trt}_{i,t-1} & \leq PS_{it} + st_j + CI_{ij,t-1}(pt_j/a_{ij}) \quad \forall (i, j) \in L, \forall t \in T \setminus \{1\} \\
(t-1)\text{per} & \leq PS_{it} \quad \forall i \in I, \forall t \in T \\
PS_{it} + st_i + pt_i Q_{it} & \leq t \text{per} \quad \forall i \in I, \forall t \in T \\
\sum_{i \in RS_j} (cp_i Q_{it} + cs_i V_{it}) & \leq \text{cap}_{jt} \quad \forall j \in RS, \forall t \in T \\
Q_{it} & \leq MV_{it} \quad \forall i \in I, \forall t \in T \\
Q_{it} & \geq 0, F I_{it} \geq 0, PS_{it} \geq 0, V_{it} \in \{0,1\}, \quad \forall i \in I, \forall t \in T \\
SL_{ijt} & \geq 0, CI_{ijt} \geq 0, \quad \forall (i, j) \in L, \forall t \in T
\end{align*}
\]
Our production planning model has two important features that distinguish it from other models that exist in the literature. These two features also make it possible for us to link the production and distribution models. First, it involves load size, component inventory and regular inventory variables. Load size variables are for the amount of an item sent from one location to another in a period. Component inventory variables represent the inventory of a particular item at a production location to be used in the production of another item which is produced at this location. Regular inventory is the inventory of an item that is held at the production location of the item. In almost all MIMLP models, the load size and the component inventory variables do not exist and all the inventory of an item is represented using the regular inventory variables, assuming that there will be instantaneous deliveries from the inventory of items at the beginning of each period if needed. As we stated earlier we do not make this assumption and therefore we had to make the distinction about the inventory of an item in the system that we described and involve the load size variables. These additions make our production model much more realistic than existing models.

The second important feature of our model is the fact that it takes the transportation time into account as a parameter in the model. The model uses two different time scales. One of these is the regular planning period, such as week or month, for which we decide the production amounts and load sizes to send to possibly multiple locations for each item at each period. The other time scale is the time unit used to represent the production and transportation times. This time unit could be minute or hour or something else depending on the problem. Hereafter, we will call this time unit the time bucket.

In the objective function of the model we include inventory holding and setup costs. We assume that the production cost does not change over time for the sake of simplicity. This is not a restrictive assumption at all since it only takes an addition of a production cost expression to the objective to relax this assumption and this addition would not affect the complexity of the model. We assign different per unit inventory holding cost to $I_{i,t}$ and $CI_{i,j,t}$ since we consider the transportation of goods as a kind of value adding operation and, consequently, per unit inventory holding cost for $CI_{i,j,t}$’s will be somewhat higher.

Constraint set (2) is the regular inventory balance constraint. If an item is an end item the expression involving the variable $SL_{i,j,t}$ in this constraint set will drop off the equation since there will be no upper layer to send the item. Instead demand of the end item will be directly satisfied. If the item is not an end item then the demand expression will drop off
the equation since we assume there is no external demand for the component items. The summation involving $S_{i,j}$ represents the total amount of component sent to one upper layer in the product structure. Constraint set (4) represents the inventory balance restriction for component inventories. The initial conditions regarding the constraint sets (2) and (4) are taken care of by constraint sets (3) and (5) respectively. Constraint set (6) states that the new load of component items must arrive before the current component inventories are consumed by the production. An assumption regarding this constraint is that the time at which we send a load of component items associated with period $t$ is the start of the period $t + 1$ which is represented by $t_{per}$ in the constraint. The expression involving $CI_{i,j-1}$ in this constraint yields the time period that it takes the production to consume the component inventory coming from previous period. Constraint set (7) simply ensures that the start time of production in a period must be greater than the start time of the period. Constraint set 8 is to make sure that production in a period ends before the end of the period. The capacity restriction is enforced by constraint set (9). Constraint set (10) is to set the binary setup variable to one when the associated amount of production variable takes a value greater than zero.

In this model, the production amount in a period is limited in three ways. First, it is restricted by the availability of subassemblies in terms of arrival time and quantity. Second, production must take place within the start-end time limits of a period. Third, total production associated with a resource should not exceed the available capacity of the resource. Here, it is implicitly assumed that a resource can process multiple items simultaneously provided that the capacity restriction on total production across items is not violated. In this formulation we considered the travel times between locations, $t_{r_{ie,je}}$, as parameters to the model. As we discussed in the problem description section, we will explain the model that takes these travel times as variables in section 2.4.

### 2.3 Transportation planning model

As a base model for the transportation planning problem, we present the formulation given in Dumas et al. (1991) for the multiple vehicle pickup and delivery problem with time windows. We closely follow the explanations in this paper as we describe the model. This formulation assumes a single depot and a homogenous fleet of vehicles.

We explain the notation and the way that the model is constructed together. Given that the problem involves $n$ requests, let node $i$ and node $n + i$ represent the pickup and delivery
location of the \(i^{th}\) request, respectively, in a network. In this network, different nodes obviously may represent the same physical location. Adding node zero and node \(2n+1\), our network has the node set \(N = \{0, 1, 2, .., n, n+1, n+2, .., 2n, 2n+1\}\). Pickup and delivery nodes are included in the sets \(P^+\) and \(P^-\) respectively where \(P^+ = \{1, 2, .., n\}\) and \(P^- = \{n+1, n+2, .., 2n\}\). \(P = P^+ \cup P^-\) is the set of the nodes other than the depot nodes.

Request \(i\) requires that \(d_i\) units be shipped from node \(i\) to node \(n+i\) and pickup and delivery must be within the time windows \([a_i, b_i]\) and \([a_{n+i}, b_{n+i}]\) respectively. Let also \([a_0, b_0]\) and \([a_{2n+1}, b_{2n+1}]\) denote the departure time and the arrival time back to the depot respectively. \(V = \{1, 2, .., |V|\}\) represents the set of vehicles indexed by \(v\). For each pair of \(i, j\) in \(N\), let \(tij\) denote the travel time and \(cij\) travel cost associated with going from the physical location of node \(i\), \(fc_i\), to physical location of node \(j\), \(fc_j\). \(s_i\) is the service time at node \(i\).

There are three variables used to construct the model: binary flow variables \(X_{ij}^v\), \(v \in V\), \(i, j \in N\), \(i \neq j\), the time variables \(TS_i\), \(i \in P\), and \(TS_0^v, TS_{2n+1}^v\), \(v \in V\) and the load variables \(Y_i\), \(i \in P\). \(X_{ij}^v\) is equal to 1 if vehicle \(v\) goes from node \(i\) to node \(j\) and equal to zero otherwise. \(TS_i\) is the time at which service at node \(i\) starts. \(TS_0^v\) and \(TS_{2n+1}^v\) represent the times at which vehicle \(v\) leaves and comes back to the depot respectively. Total load on a vehicle right after it leaves a node is represented by \(Y_i\). At the start, vehicles are assumed to be empty, i.e. \(Y_0 = 0\).

There is a cost function \(g(Y_i)\) in the objective which might be excluded from the formulation depending on the problem characteristics. \(g(Y_i)\) denotes a non-decreasing function of the total load on a vehicle just after it leaves node \(i\), \(i \in P\). This function takes the load factor into consideration in calculation of the travel cost.

Summary of the notation:

\(X_{ij}^v\) : binary flow variables which are equal to 1 if vehicle \(v\) goes from node \(i\) to node \(j\) and equal to zero otherwise
\(TS_i\) : the time at which service at node \(i\) starts
\(TS_0^v\) and \(TS_{2n+1}^v\) : the times at which vehicle \(v\) leaves and comes back to the depot respectively
\(Y_i\) : total load on a vehicle right after it leaves a node \(i\)
$d_i$: number of units to be shipped from node $i$ to node $n + i$

$s_i$: the service time at node $i$

$f_{c_i}$: physical location of node $i$

$[a_i, b_i]$: service time window for node $i$

$trt_{f_{c_i}, f_{c_j}}$: the travel time associated with going from the physical location of node $i$, $f_{c_i}$, to physical location of node $j$, $f_{c_j}$

$c_{f_{c_i}, f_{c_j}}$: the travel cost associated with going from the physical location of node $i$, $f_{c_i}$, to physical location of node $j$, $f_{c_j}$

The pickup and delivery model will serve as the transportation model in the remainder of this report. Following is the formulation of the pickup and delivery problem with time windows.

$$
\text{Min} \sum_{u \in V} \sum_{i \in N} \sum_{j \in N} g(Y_i) c_{ij} X_{ij}^v
$$ (13)

Subject to

$$\sum_{u \in V} \sum_{j \in N} X_{ij}^v = 1, \quad i \in P^+$$ (14)

$$\sum_{j \in N} X_{ij}^v - \sum_{j \in N} X_{ji}^v = 0, \quad i \in P, v \in V$$ (15)

$$\sum_{j \in P^+} X_{ij}^v = 1, \quad v \in V$$ (16)

$$\sum_{i \in P^-} X_{i,2n+1}^v = 1, \quad v \in V$$ (17)

$$\sum_{j \in N} X_{ij}^v - \sum_{j \in N} X_{ji}^v = 0, \quad i \in P^+, v \in V$$ (18)

$$TS_i + s_i + trt_{f_{c_i}, f_{c_i+n}} \leq TS_{n+i}, \quad i \in P^+$$ (19)

$$X_{ij}^v = 1 \Rightarrow TS_i + s_i + trt_{f_{c_i}, f_{c_j}} \leq TS_j, \quad i, j \in P, v \in V$$ (20)

$$X_{0j}^v = 1 \Rightarrow TS_0^v + trt_{f_{c_0}, f_{c_j}} \leq TS_j, \quad j \in P^+, v \in V$$ (21)
\begin{align*}
X_{i,2n+1}^v = 1 & \Rightarrow TS_i + s_i + \text{trt}_{j,e_i \in 2n+1} \leq TS_{2n+1}^v, \\
i \in P^-, v \in V \quad (22) \\
a \leq TS_i \leq b, \\
i \in P \quad (23) \\
a_0 \leq TS_0^v \leq b_0, \\
v \in V \quad (24) \\
a_{2n+1} \leq TS_{2n+1}^v \leq b_{2n+1}, \\
v \in V \quad (25) \\
X_{ij}^v = 1 & \Rightarrow Y_i + d_i = Y_j, \\
i \in P, j \in P^+, v \in V \quad (26) \\
X_{ij}^v = 1 & \Rightarrow Y_i - d_{j,n} = Y_j, \\
i \in P, j \in P^-, v \in V \quad (27) \\
X_{0j}^v = 1 & \Rightarrow Y_0 + d_j = Y_j, \\
j \in P^+, v \in V \quad (28) \\
d_i \leq Y_i \leq D, \\
i \in P^+ \quad (29) \\
Y_0 = 0 \\
(30) \\
X_{ij}^v \in \{0, 1\}, \\
i, j \in N, v \in V \quad (31)
\end{align*}

The objective function of the formulation represents the total travel cost of the vehicles. Constraint set (14) enforces that each pickup node is visited once by one of the vehicles. Constraint set (15) ensures that if a vehicle enters a node it must exit it. The first visit by all vehicles must be to a pickup node and the last visit must be to a delivery node. These two conditions are stated in constraint sets (16) and (17) respectively. Constraint set (18) makes sure that if a vehicle visits a pickup node then it must visit the associated delivery node. Pickups must come before deliveries. This is enforced by constraint set (19). Constraint sets (20)-(22) describe the compatibility requirements between routes and schedules, while constraint sets (23)-(25) are the time windows constraints. Constraint sets (26)-(28) express the compatibility requirements between routes and vehicle loads, while constraint set (29) is the capacity constraint.

Any route in a solution to the model will be elementary, i.e. without cycles due to constraint sets (20)-(22) since these constraints enforce increasing service times, $TS_i$'s, at the nodes of the route. As a special structure in the formulation, constraint sets (14)-(17) and (29) form a multi commodity flow problem. Another special structure is that, for a
given vehicle, constraint sets (15)-(25) describe a path starting and ending at the depot under pairing constraints (18), precedence constraints (19), time window constraints (20)-(25) and capacity constraints (26)-(30). That is, this structure represents a constrained shortest path problem. Dumas et al. (1991) utilize this structure in their algorithm to find optimal solutions.

2.4 Linkage between the two models

There are three variables/parameters that link the production planning and the transportation planning models. These are load sizes, $S_{L_{i,j}}$'s, travel time allowances, for which we give notation later, and service time windows, $[a_i, b_i]$'s. In our perception of the problem, we will assume that the service time windows are fixed as parameters in both models and consider only load sizes and the travel time allowances as the variables that link two models. In this section, we will modify both models so that they will determine the linking variables independently. In other words, these modified models will give solutions that are best in terms of their own objectives. We will describe an approach in Section 3 to simultaneously determine the linking variables which optimize the total cost over both models.

We can solve the production and transportation planning problems in a sequential fashion. We first solve the production planning model and obtain load sizes and travel time allowances. Using these load sizes and travel time allowances we solve the transportation model. As a result we get a complete solution consisting of production and transportation plans. This solution is dominated by the production planning part. It will provide an upper bound to the optimal solution of the integrated model since it is just a feasible solution to the integrated model. There might exist another solution which involves much less transportation cost and not much higher production cost compared to sequential solution and is better in terms of total cost. For a solution to be optimal in our problem setting, it needs to be determined by considering the cost factors of both models simultaneously.

2.4.1 Modifications to the production planning model

In order to build the linkage between the two models, we had to make some modifications to them. For the production planning part, one of the modifications necessary will be in constraint set (6). Let $R_{l,c_i|c_j}$ denote the travel time allowance variable for going from $l_{c_i}$,
the production location for product \( i \), to \( l_{c_j} \), the production location for product \( j \), and \( tw \) the time window length parameter. \( R_{l_{i_{c_j}}} \) is a multiplication factor representing the travel time allowance in terms of the number of direct travel time periods. Actual travel time allowance in time buckets is given by the term \( tr_{l_{i_{c_j}}} \cdot R_{l_{i_{c_j}}} \). We assume that \( tw \) is the same for all the pickups and deliveries in the problem in order to simplify the illustrations. We replace constraint set (6) in the production model with the following constraint:

\[
(t - 1)per + 2tw + tr_{l_{i_{c_j}}} \cdot R_{l_{i_{c_j}}} \leq PS_{jt} + st_j + CI_{i_{j-1}}(pt_j/a_{ij})
\]

\[\forall(i, j) \in L, \forall t \in T \setminus \{1\}\] (32)

In this constraint, total travel time for the items includes travel time allowance and both pickup and delivery time windows. The implicit assumption here is that we make production planning based on the worst case scenario where the items are always delivered to one upper layer at the end of the delivery service time window. Hence, total travel time is considered to be \( 2tw + tr_{l_{i_{c_j}}} \cdot R_{l_{i_{c_j}}} \) in the constraint.

Because of the assumption we made about the sending time of the loads in the production model we fix pickup time windows to be at the beginning of the periods. Delivery time windows will be dependent on the value that \( R_{l_{i_{c_j}}} \) will take. We also add the following constraint to put lower and upper bounds on \( R_{l_{i_{c_j}}} \):

\[
lub_{lb} \leq R_{l_{i_{c_j}}} \leq lub_{ub}
\]

\[\forall(i, j) \in L\] (33)

where \( lub_{lb} \) and \( lub_{ub} \) represent the lower and upper bounds, respectively.

2.4.2 Modifications to the transportation planning model

We will continue with the modifications we need to make to the transportation model in this section. In the production model, load size variables have three indices, \( i, j, t \). Since each load corresponds to a request in the transportation model and each pickup is represented by a single index, we need to convert each triply indexed set \( i, j, t \) to a single pickup node index. In order to do this, we first define predecessor sets for the items as follows:

\( PR_j : \) Set of items needed to produce item \( j \) .
That is, each \( i, j, t \) triple, \( j \in I, i \in PR_j, t \in T \), represents a load variable. The following expression converts the three indices to a unique single pickup node index:

\[
i, j, t \cdot l = \sum_{k=1}^{j-1} |PR_k|(|T| - 1) + (ord(i) - 1)(|T| - 1) + t
\]  

(34)

where \( l \) is the pickup node index. For a given \( i, j, t \), the term \( \sum_{k=1}^{j-1} |PR_k|(|T| - 1) \) gives the number of the pickup node indices used to represent the load size variables until the previous item, \( j - 1 \). There will be \(|T| - 1\) unique pickup node indices in the transportation model for each pair of \( i, j \). We use \(|T| - 1\) here since there will be no load send in the last period. Therefore, the total number of indices in the transportation model for load variables involving the predecessor items of the items 1, 2, 3, ..., \( j - 1 \) will be \( \sum_{k=1}^{i-1} |PR_k|(|T| - 1) \). \( ord(i) \) in this expression is the rank of item \( i \) in the predecessor set that item \( i \) belongs to. There will be \((ord(i) - 1)(|T| - 1)\) indices used for load variables associated with the predecessors of item \( j \) that come before item \( i \) in the set \( PR_j \). Finally, \( t \) will distinguish the load variables that have the same \( i, j \) indices. We illustrate this conversion in the example that follows.

**Conversion example** : Let us assume that the following is the product structure of a supply chain and the length of the planning horizon, \(|T| = 3\).

![Product Structure Diagram](image)

Predecessor sets of this structure are \( PR_1 = \{2, 3\}, \ PR_2 = \{4, 5\}, \ PR_3 = \{4, 5\} \). There are six links in this example and each of them corresponds to an item-predecessor pair. In total, we need to have \( 6(|T| - 1) = 6 \times 2 = 12 \) pickup node indices in the transportation model to represent each load variable. Following are the conversions of this example:
\[ i, j, t \quad \bullet \quad l \]

\[ 2, 1, 1 \quad \sum_{k=1}^{l-1} |PR_k|(3 - 1) + (\text{ord}(2) - 1)(3 - 1) + 1 = 0 + (1 - 1)x2 + 1 = 1 \]

\[ 2, 1, 2 \quad \sum_{k=1}^{l-1} |PR_k|(3 - 1) + (\text{ord}(2) - 1)(3 - 1) + 2 = 0 + (2 - 1)x2 + 2 = 2 \]

\[ 3, 1, 1 \quad \sum_{k=1}^{l-1} |PR_k|(3 - 1) + (\text{ord}(3) - 1)(3 - 1) + 1 = 0 + (2 - 1)x2 + 1 = 3 \]

\[ 3, 1, 2 \quad \sum_{k=1}^{l-1} |PR_k|(3 - 1) + (\text{ord}(3) - 1)(3 - 1) + 2 = 0 + (2 - 1)x2 + 2 = 4 \]

\[ 4, 2, 1 \quad \sum_{k=1}^{2l-1} |PR_k|(3 - 1) + (\text{ord}(4) - 1)(3 - 1) + 1 = 2x2 + (1 - 1)x2 + 1 = 5 \]

\[ 4, 2, 2 \quad \sum_{k=1}^{2l-1} |PR_k|(3 - 1) + (\text{ord}(4) - 1)(3 - 1) + 2 = 2x2 + (1 - 1)x2 + 2 = 6 \]

\[ 5, 2, 1 \quad \sum_{k=1}^{2l-1} |PR_k|(3 - 1) + (\text{ord}(5) - 1)(3 - 1) + 1 = 2x2 + (2 - 1)x2 + 1 = 7 \]

\[ 5, 2, 2 \quad \sum_{k=1}^{2l-1} |PR_k|(3 - 1) + (\text{ord}(5) - 1)(3 - 1) + 2 = 2x2 + (2 - 1)x2 + 2 = 8 \]

\[ 4, 3, 1 \quad \sum_{k=1}^{3l-1} |PR_k|(3 - 1) + (\text{ord}(4) - 1)(3 - 1) + 1 = (2x2 + 2x2) + (1 - 1)x2 + 1 = 9 \]

\[ 4, 3, 2 \quad \sum_{k=1}^{3l-1} |PR_k|(3 - 1) + (\text{ord}(4) - 1)(3 - 1) + 2 = (2x2 + 2x2) + (1 - 1)x2 + 2 = 10 \]

\[ 5, 3, 1 \quad \sum_{k=1}^{3l-1} |PR_k|(3 - 1) + (\text{ord}(5) - 1)(3 - 1) + 1 = (2x2 + 2x2) + (2 - 1)x2 + 1 = 11 \]
In this way we obtain a unique pickup node index for each load variable. We will use this conversion procedure in our modifications to the transportation model. We first add the following constraint for number of units to be shipped:

\[ d_i = SL_{i,j,t} \quad \forall (i,j) \in L, \forall t \in T \setminus \{T\} \]  

(35)

\( l \) is as we defined in equation (34). Next we add several sets of constraints involving and restricting \( SL_{i,j,t} \)'s, the load size variables. These constraints are analogous to flow conservation constraints in network models and basically state the fact that inputs to an item node in the product structure from each incoming arc by time period \( t \) must be greater than or equal to the total output across all the outgoing arcs by period \( t + 1 \). This one period gap between input and output is due to the assumption that any load sent in period \( t \) will arrive at the upper layer in period \( t + 1 \). Let \( E \) denote the set of end items. We add the following constraint sets to the transportation model for end items:

\[ in\!c\!i_{ij} + \sum_{l=1}^{T_i-1} SL_{ij,l} \geq a_{ij} \left( \sum_{l=1}^{T_i} \text{dem}_{jl} - \text{ini}_j \right) \quad \forall j \in E, \forall i \in PR_j, \forall T_0 \in T \setminus \{1\} \]  

(36)

\[ in\!c\!i_{ij} \geq a_{ij} (\text{dem}_{jl} - \text{ini}_j) \quad \forall j \in E, \forall i \in PR_j \]  

(37)

Let \( B \) denote the items that are neither end item nor the lowest layer item in the product structure. For these middle level items, we replace the demand term in equations (36) and (37) with the total amount of load that is sent from the item node to all successor item nodes in the upper layer.

\[ in\!c\!i_{ij} + \sum_{l=1}^{T_i-1} SL_{ij,l} \geq a_{ij} \left( \sum_{k=1}^{T_i} \sum_{k \in N_j} SL_{jkd} - \text{ini}_j \right) \quad \forall j \in B, \forall i \in PR_j, \forall T_0 \in T \setminus \{1\} \]  

(38)

\[ in\!c\!i_{ij} \geq a_{ij} \left( \sum_{k \in N_j} SL_{jkd} - \text{ini}_j \right) \quad \forall j \in B, \forall i \in PR_j \]  

(39)
Since we assume there is an infinite supply of subassemblies for the lowest level items we will not add similar constraints for these items. By adding constraint sets (35), (36), (37), (38), and (39), we make sure that the load sizes, $d_i$ or $SL_{i,j}$'s, which are determined by the transportation model with no regard to the production related costs, satisfy the demand of end item and obey the restriction of the conservation of item flows we described above. We can see here that some of the load variables can take zero value in a solution. Since only the pickup points whose associated load variable has positive value need to be visited by a vehicle, we replace constraint set (14) with the following constraint:

$$M \left[ \sum_{v \in V} \sum_{j \in N} X_{ij}^v \right] \geq d_i, \quad i \in P^+$$  \hspace{1cm} (40)

Where $M$ is a large number. Another modification we need to make to the transportation model concerns the time windows. We had both pickup and delivery time windows fixed as parameters in our previous transportation model. In the modified model, we will also have pickup time windows fixed as parameters but the delivery time windows will depend on the travel time allowances variable. We represented this variable as $R_{i,c,l,c}$ in the modified production model. First we define some parameters for the modified transportation model. Let $a_l = t_{per}$ where $t$ comes from the triple $i, j, t$ which is used to find $l$. $a_l$ designates the start time of the pickup time window for pickup node $l$ and this time is basically the end of the period that the load variable $SL_{i,j}$ is associated with. Let also $f_{c_l} = l_{c_l}$ and $f_{c_{l+n}} = l_{c_j}$ where again $i, j$ comes from the triple $i, j, t$ which is used to find $l$. Here $n$ represents the total number of total requests. $f_{c_l}$ is the pickup location associated with the pickup node $l$ and $f_{c_{l+n}}$ is the delivery location of this pickup which corresponds to the delivery node $l + n$. These pickup and delivery locations are the production locations of items $i, j$, i.e. $l_{c_l}, l_{c_j}$. We replace the constraint set (23) in the transportation model with the following ones:

$$a_l \leq TS_i \leq a_l + tw \quad l \in P^+ \quad (41)$$

$$a_{i-n} + tw + R_{j,c_{i-n},j} trt_{j,c_{i-n},j} \leq TS_i \leq a_{i-n} + 2tw + R_{j,c_{i-n},j} \quad i \in P^- \quad (42)$$

Constraint set (41) imposes the pickup time window restriction which is considered to be a fixed parameter. Constraint set (42) describes the delivery time window in terms of
travel time allowance variable $R_{fe_i}$, As we did for the modified production model, we put upper and lower bounds on $R_{fe_i}$ by adding the following constraint;

$$l_{urb} \leq R_{fe_i} \leq u_{urb}$$

$i,j \in N$  \hspace{1cm} (43)

After all these modifications, we finally have a transportation model which involves the load sizes and the travel time allowances as variables, like the modified production model, in contrast to our previous transportation model which considers these two variables to be fixed parameters. In section 3.2 we will explain a Lagrangean decomposition approach that utilizes the modified production and transportation models to make integrated planning decisions.

3. The value of integrating production and transportation planning

In this section we will demonstrate the value of integrating production and transportation decisions using small size examples. Because a model including both production and transportation models will be too large to handle we will apply Lagrangean decomposition to make integrated decisions using two models.

As we explained in the previous section, there are two sets of variables, namely load size and travel time allowance variables, that link the production and transportation models. First we will consider the case where the travel time allowance variables are fixed and try to make integrated decisions about load variables. Then, we will consider the case where the load variables are fixed by the production model and the integrated decisions involve only the time allowance variables. That is, we will deal with the linking variables separately in order to see the effects of the set of variables separately, and to avoid the computational burden that would come if we took into account both sets of variables simultaneously.

3.1 Lagrangean Decomposition

We give a brief explanation of Lagrangean decomposition. Lagrangean decomposition was first proposed by Guignard and Kim (1987a, b). The technique has been applied to several discrete optimization problems (Guignard and Rosewein (1989), Millar and Yang (1992), Fumero and Vercellis (1997)). It assumes that a MIP problem has a constraint set which
consists of several special structures. It is possible to define a set of subproblems, one for each special structure in the original constraint set. Each subproblem uses a "copy" of the original variable set. The constraint that the copies of the variable have to be equal in the final solution, called the coupling constraint, is Lagrangean relaxed and reflected in the objectives of the subproblems. Unlike Lagrangean relaxation which removes a part of the constraint set, Lagrangean decomposition retains all of the original constraints. It has been proven that Lagrangean decomposition produces bounds that are at least as good and can be better than the bounds from lagragean relaxation (Guignard and Kim (1987a, b)).

3.2 Lagrangean decomposition of the integrated model

In our problem, the integrated model involves two special structures, production and transportation models. Our aim is to decompose this large model into production and transportation subproblems. The production and transportation models we mention here are the modified models that we explained in previous section. We define the following notation:

\[ P : \text{integrated model} \]
\[ P_P : \text{production model} \]
\[ P_T : \text{transportation model} \]
\[ f_P : \text{Objective function of the production model} \]
\[ f_T : \text{Objective function of the transportation model} \]
\[ C_P : \text{constraint set of the production model} \]
\[ C_T : \text{constraint set of the transportation model} \]
\[ SL1_{ijt} : \text{load variable of the production model (Amount of item} i \text{ sent for the production of item} j \text{ in period} t) \]
\[ SL2_{ijt} : \text{load variable of the transportation model} \]
\[ SL1 : \text{vector of the load variables of the production model} \]
\[ SL2 : \text{vector of the load variables of the transportation model} \]
\[ R1_{ij} : \text{travel time allowance variable of the production model (Number of direct travel time periods allowed to carry goods from location} i \text{ to location} j) \]
\[ R2_{ij} : \text{travel time allowance variable of the transportation model} \]
\[ R1 : \text{vector of the travel time allowance variables of the production model} \]
\[ R2 : \text{vector of the travel time allowance variables of the transportation model} \]
\[ \lambda_{SL} : \text{vector of multipliers for the load variables} \]
\[ \lambda_R : \text{vector of multipliers for the travel time allowance variables} \]
\( v(\cdot) \): optimal value for problem \((\cdot)\)

We define problem \( P \) as follows:

\[
P : \min f_p + f_T \\
subject to \\
C_p, C_T, \\
SL_{1ij} = SL_{2ij}, \quad \forall i, j, t \\
R_{1ij} = R_{2ij}, \quad \forall i, j, t
\]

The last two equalities are the coupling constraints to which we apply Lagrangean relaxation to get the Lagrangean decomposed model, i.e. \((LD)_{\lambda_L, \lambda_R}\), as follows:

\[
(LD)_{\lambda_L, \lambda_R} : \text{Min} \{ f_p + f_T + \lambda_L (SL_2 - SL_1) + \lambda_R (R_2 - R_1) \text{ s.t. } C_p, C_T \}
\]

\[
= \left\{ \text{Min} \{ f_p - \lambda_L SL_1 - \lambda_R R_1 \text{ s.t. } C_p \} + \text{Min} \{ f_T + \lambda_L SL_2 + \lambda_R R_2 \text{ s.t. } C_T \} \right\}
\]

We define the following problems:

\[
P_{PP} : \text{Min} \{ f_p - \lambda_L SL_1 - \lambda_R R_1 | C_p \}
\]

\[
P_{TT} : \text{Min} \{ f_T + \lambda_L SL_2 + \lambda_R R_2 | C_T \}
\]

The Lagrangean decomposition dual problem \((LDD)\) is defined as follows:

\[
LDD : \text{Max} \{ (LD)_{\lambda_L, \lambda_R} \}
\]

### 3.3 Subgradient search algorithm to solve the Lagrangean decomposition dual

We present a subgradient search algorithm to solve \(LDD\) and to get upper bounds as the iterations progress. We define the following additional notation to describe the algorithm:

\(UB^j\) : Upper bound found in iteration \(j\)

\(LB^j\) : Lower bound found in iteration \(j\)

\(UB^*\) : Current best upper bound
\( \gamma^j \): Scale factor at iteration \( j \). It is set to 2 initially and reduced by half if the lower bound does not improve in four iterations.

\( u^j \): Step size of the subgradient search.

The subgradient search algorithm we apply is as follows:

Step1: Set \( j = 1 \). Initialize \( \lambda_{SL}, \lambda_{R}, \gamma^j \), and \( UB^* \).

Step2: Solve \( P_{PP} \) and \( P_{TT} \). Set \( LB^j = u(P_{PP}) + u(P_{TT}) \).

Step3: Determine \( UB^j \). Set \( UB^* = \min\{UB^j, UB^*\} \).

Step4: Determine the duality gap as follows:

\[
\theta = \frac{2(UB^* - LB^j)}{UB^* + LB^j} \times 100
\]

Stop if \( j > \text{iteration limit} \) or \( \theta < \epsilon \) where \( \epsilon \) is the prespecified tolerance.

Otherwise go to step 5.

Step5: Update multipliers \( \lambda_{SL}, \lambda_{R} \) and step size \( u^j \). Set \( j = j + 1 \) and go to step 2.

\( \theta \) in step 4 is the ratio of the difference between current best upper bound and the current lower bound to the average of these two values.

As we stated earlier, we will apply Lagrangean decomposition on two sets of linking variables separately. Step 3 and step 5 of the subgradient search algorithm will differ depending on which set of variables we consider.

**Load variable case:**

In this case we basically fix the \( R1 \) and \( R2 \) variables (defined in section 3.2) to a prespecified value and set the multipliers in \( \lambda_{R} \) to zero during entire iterations. We update the multipliers \( \lambda_{SL} \) and step size in step 5 as follows:

\[
\lambda_{SL}^{j+1} = \lambda_{SL}^j + u^j(SL2 - SL1)
\]

where \( u^j \) is given by:
\[ u^j = \frac{\gamma^j (UB^* - LB^j)}{\sum_i \sum_j (SL_{1ij} - SL_{2ij})^2} \]

We find the upper bound in step 3 as follows. After solving \( P_{PP} \) and \( P_{TT} \) to get \( LB^j \), we take the values of \( SL_{1ij} \)'s from the solution of \( P_{PP} \) and fix the values of load variables to these values both in \( P_P \) and \( P_T \) and solve these problems. The upper bound, \( UB^j \), is:

\[ UB^j = v(P_P) + v(P_T) \]

where the load variables are fixed as described.

**Travel time allowance variable case:**

First, we solve the production model \( P_p \) using average values for \( R1 \) variables to obtain a solution. The average values for \( R1 \) are the average of upper and lower bounds of \( R1 \). The values of the load size variables both in \( P_{PP} \) and \( P_{TT} \) are set to the values obtained from this first solution of \( P_P \) during all iterations. The multipliers \( \lambda_{SL} \) are set to zero during entire iterations.

We update the multipliers and step size in a similar manner to the previous case as follows:

\[ \lambda^{j+1}_R = \lambda^j_R + u^j (R2 - R1) \]

where \( u^j \) is given by:

\[ u^j = \frac{\gamma^j (UB^* - LB^j)}{\sum_i \sum_j (R1_{ij} - R2_{ij})^2} \]

We find the upper bound in step 3 in the following way. After solving \( P_{PP} \) and \( P_{TT} \) to get \( LB^j \), we take the values of \( R2_{ij} \)'s from the solution of \( P_{TT} \) and set the travel time allowance variables both in \( P_P \) and \( P_T \) to these values and solve these problems. The upper bound, \( UB^j \), is given by:

\[ UB^j = v(P_P) + v(P_T) \]
where the travel time allowance variables are fixed as described.

3.4 Computational experiments

We apply the Lagrangean decomposition scheme on a small size problem with four items that are produced in four locations. The planning period is also set to four. We have used the AMPL modeling environment and CPLEX solver to implement the Lagrangean decomposition scheme. The product structure of the problem is depicted in the following figure:

![Product Structure Diagram]

We assume that there are two vehicles available to provide transportation to the system. There is also a depot location from which the vehicles are dispatched. The production period is assumed to be on the order of weeks while the production and transportation time units are assumed to be in minutes.

Other parameters of the problem, such as inventory holding and setup costs, distances between locations, and travel time between locations, are determined in an arbitrary manner using the random number generator of MS Excel. Our experiments are divided into two parts: 1) Those that involve load variables and 2) Those that involve travel time allowance variables.

The results we obtained for the load variable case are summarized in Table 1.

Table 1. Computational Results Using Load Variables as the Linkage
Problem type II is exactly the same as problem type I except the setup cost in problem type II is 1/4 of those in problem type I and the unit travel time in problem type II is 1/2 of those in problem type I. Unit travel time is the time to travel one unit of distance. We expected that problem type II will have more balanced final solutions than the case where the production or transportation part dominates the final solutions. Type II problems required more computational effort than type I. We solved the same problems (either problem type I or II) using different unit travel costs. We set the iteration limit to 40 and quit the iterations if the optimal solution is found. The sequential solutions in column 5 are the ones obtained using the sequential approach where we first solve the production model and then the transportation model by taking the load variables of the production solution fixed. These solutions are taken as the initial upper bound in the subgradient search. % change in the production and transportation plans are given in columns 3 and 4. This percentage is defined as follows:

\[
\text{% change} = 100 \times \left\{ \frac{\sum |\text{independent solutions} - \text{final solutions}|}{\sum |\text{independent solutions}|} \right\}
\]

The independent solutions are the load variable values in the solutions obtained by both models at the first iteration of the subgradient search when the multipliers are all zero. That is, these are the solutions that are best in terms of respective objectives of the two models with no regard to the objective of the other model. The final solutions are the load variables in the solution corresponding to the best upper bound found by subgradient search. The table also includes best lower bounds, % gap between best upper bound and best lower bounds, which is denoted by $\theta$ in the subgradient search algorithm above, and the % gain of the best upper bound over the sequential solution. To reduce the computational burden, we have given the solution found by the submodels in the previous iteration as the starting solution at each iteration.
As expected, we observe in % change columns that as the unit transportation cost increases the best upper bound solution, or integrated solution in other words, becomes the solution dictated by the transportation model and that type II problems gave more compromised solutions than type I problems. % gain or the value of integration also increases as the unit transportation cost increases, which is a trend one could anticipate.

We used only problem type I for travel time variable case. Table 2 summarizes the solutions we obtained.

Table 2. Computational Results Using Travel Time Variables as the Linkage

<table>
<thead>
<tr>
<th>Unt. Trv. Cost</th>
<th>% change in Pro. Pl.</th>
<th>% change in Trn. Pl.</th>
<th>Sequential solution</th>
<th>Best UB</th>
<th>Best LB</th>
<th>% Gap</th>
<th>% Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>3.11</td>
<td>2.775</td>
<td>131486</td>
<td>131271</td>
<td>131271</td>
<td>0</td>
<td>.16</td>
</tr>
<tr>
<td>25</td>
<td>7.96</td>
<td>1.79</td>
<td>137606</td>
<td>136531</td>
<td>136531</td>
<td>0</td>
<td>.78</td>
</tr>
<tr>
<td>50</td>
<td>7.96</td>
<td>1.79</td>
<td>142256</td>
<td>143106</td>
<td>143106</td>
<td>0</td>
<td>1.48</td>
</tr>
<tr>
<td>100</td>
<td>9.85</td>
<td>3.58</td>
<td>1602056</td>
<td>156256</td>
<td>156256</td>
<td>0</td>
<td>2.68</td>
</tr>
<tr>
<td>200</td>
<td>9.85</td>
<td>3.58</td>
<td>191156</td>
<td>182556</td>
<td>182556</td>
<td>0</td>
<td>4.50</td>
</tr>
</tbody>
</table>

These problems took far less computation time than the problems involving load variables. Optimal solutions are found in three iterations for all the problems. Definitions of the columns are the same as before. In this table, we see similar trends to the load variables experiments.

In order to see the effects of different distribution patterns of the locations of facilities (5 facilities in total including the depot) using the load variable as the linkage between the two models, we have generated 8 random problems. In all of these problems production part of the problem is the same as the ones used in previous experiments. We have set the unit travel cost to 3 and unit travel time to 5 for all problems after some trial and error experiments to produce feasible and reasonable problems in terms of relative cost of the transportation. The transportation cost, on the average, turned out to be 10-20% of the total cost out of 8 problems. The problems differ only in terms of the location of production facilities and the depot. These locations are distributed uniformly over a 30x30 square for each problem. The results are summarized in Table 3.

Table 3. Effects of the distribution of locations using Load Variables as the Linkage
Problem number | Average distance | Sequential Solution | Best UB | Best LB | % Gap | % Gain |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11.5</td>
<td>34266</td>
<td>33996</td>
<td>33989</td>
<td>0.02</td>
<td>0.79</td>
</tr>
<tr>
<td>2</td>
<td>11.1</td>
<td>35391</td>
<td>34801</td>
<td>33794</td>
<td>2.89</td>
<td>1.67</td>
</tr>
<tr>
<td>3</td>
<td>20.4</td>
<td>40026</td>
<td>37336</td>
<td>35853</td>
<td>3.97</td>
<td>6.72</td>
</tr>
<tr>
<td>4</td>
<td>13</td>
<td>34446</td>
<td>34446</td>
<td>34423</td>
<td>0.07</td>
<td>0.00</td>
</tr>
<tr>
<td>5</td>
<td>10.1</td>
<td>33996</td>
<td>33996</td>
<td>33990</td>
<td>0.02</td>
<td>0.00</td>
</tr>
<tr>
<td>6</td>
<td>10.3</td>
<td>34716</td>
<td>34716</td>
<td>33657</td>
<td>3.05</td>
<td>0.00</td>
</tr>
<tr>
<td>7</td>
<td>12.2</td>
<td>36381</td>
<td>35071</td>
<td>34022</td>
<td>2.99</td>
<td>3.50</td>
</tr>
<tr>
<td>8</td>
<td>18.5</td>
<td>37101</td>
<td>35926</td>
<td>34907</td>
<td>2.84</td>
<td>3.17</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.98</td>
<td>1.99</td>
</tr>
</tbody>
</table>

Second column of table 3 gives the average distance from each facility to the other facilities. The definitions for the other columns are the same as those we have used before. On this table, one can see that the benefit of the integrated solution is in general proportional to the average distance between facilities, which simply determines how important or costly the transportation part of the problem is. If we look at the % gap column, we can see that the final results are within 4% of the actual integrated optimal solution and on the average within 2% of the optimal. The % gain varies significantly depending of the structure of the transportation problem, and it is, on the average, 2%. Considering the fact that the transportation cost to total cost ratio is tried to be limited within a realistic range (10-20%), the savings due to integrated solution can amount to significant figures depending on the actual total cost. The sequential solution for problem 4 and 5 is either the integrated optimal as well or very close to it, which shows the fact that it does not always yield a gain to solve the problem as an integrated one.

Based on this small experimental study, we conclude that the saving due to integrated decision making can be substantial depending on the cost structure of the problems.

### 3.5 A sensitivity analysis on cost parameters

In order to see the effect of the three cost parameters of the problem (namely, unit inventory holding cost, setup cost, and unit travel cost) on the integrated solution we ran experiments using two levels of these factors. We set the planning horizon to four periods and assumed that each facility produces a single type item. The iteration limit of the subgradient search is fixed at 60. We set the iteration limit higher than that of previous experiments in order to see the effects of cost parameters better. There are eight problems corresponding to the combinations of the three factors. We solved two sets of eight problems, one set for the case with general product structure and the other one for serial product structure. The serial structure we use in the experiments is the following:
The general product structure we used here is the same as the one we had before in previous section. The problem data is, however, some what different than before. The results for the problems with serial and general product structures are summarized in Tables 3 and 4, respectively.

### Table 3. Computational Results for Problems with a Serial Product Structure

<table>
<thead>
<tr>
<th>Cost Parameters</th>
<th>No. Tran. Lots</th>
<th>No. Prod. Lots</th>
<th>Total Cost</th>
<th>Tran. Cst /Prod. Cst.</th>
<th>% Change</th>
<th>Gain</th>
<th>Gap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inv. Setup Tran.</td>
<td>Inv. Setup Tran.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L L L 6 10 12379 5 28800 11900 .29 140 28 5.08 7.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L L H 6 10 12379 28800 59500 1.44 48 111 12.69 9.92</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L H L 6 9 14299 137756 11900 .08 140 28 1.7 2.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L H H 6 9 14299 137756 59500 .39 51 29 6.47 8.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H L L 8 12 36300 35353 15450 .21 144 25 2.1 9.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H L H 5 12 36300 42999 59500 .75 48 111 12.58 21.98</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H H L 6 10 61897 144000 11900 .06 140 29 1.09 1.73</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H H H 6 10 61897 144000 59500 .29 51 29 5.07 7.12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 4. Computational Results for Problems with General Product Structure

<table>
<thead>
<tr>
<th>Cost Parameters</th>
<th>No. Tran. Lots</th>
<th>No. Prod. Lots</th>
<th>Total Cost</th>
<th>Tran. Cst /Prod. Cst.</th>
<th>% Change</th>
<th>Gain</th>
<th>Gap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inv. Setup Tran.</td>
<td>Inv. Setup Tran.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L L L 4 6 15306 16300 9700 .31 100 153 15.80 2.32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L L H 6 9 7606 23600 76500 2.45 100 52 10.58 35.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L H L 4 6 15306 81500 9700 .10 100 0 0 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L H H 4 6 15306 81500 48500 .50 100 0 0 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H L L 12 14 4780 34900 26000 .65 137 0 0 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H L H 9 12 14755 31900 104250 2.23 100 51 11.06 56.42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H H L 4 6 76530 81500 9700 .06 100 153 3.54 1.35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H H H 4 6 76530 81500 48500 .31 100 153 15.80 2.32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

-30-
We make the following observations based on the two result tables above;

1. There exists a consistent proportionality between transportation cost parameter and % gain. This agrees with intuition. One expects more room for improvement over the sequential solution as the transportation part of the problem becomes more important.

2. If we consider the runs LLL and HLL, we see an increase in the number of transportation lots and total transportation cost. This is explainable by the fact that we tend not to carry bulky loads which, in turn, would lead to bulky inventories. Inventory cost in run HLL is much higher than that of run LLL and hence, the result of run HLL shows small lot sizes and small transportation loads compared to run LLL.

3. % gain and % gap are pretty much proportional. This can be explained as follows. Subgradient search basically converts a sequential initial solution to an integrated one. % gain is high when the transportation cost parameter is high in general. It can be anticipated that as the transportation cost parameter gets higher the sequential initial solution deviates more from the integrated solution since the sequential solution is the solution where there is no regard to transportation cost. As the sequential initial solution deviates more from the integrated solution, % gap gets larger since % gap is simply a measure of the success of the subgradient search with limited number of iterations in converting the sequential initial solution to an integrated one.

4. There is a consistency between transportation cost to production cost ratio and % gain and again between transportation cost to production cost ratio and % gap.

5. Number of production lots is reduced on the average for the obvious reason when we increase the inventory cost parameter.

5. Conclusion

There is a lack of research and practice in approaching production and transportation planning in an integrated manner. In this paper, we suggest new production and transportation planning models which include required adjustments and additions to the
existing ones in order to simultaneously consider the cost factor of both models. We
develop a Lagrangean based approach to solve the integrated model of production and
transportation planning. It is demonstrated in our experiments that saving in total cost due
to integrated decision making could be significant depending on the cost structure of the
problems. Future research should be directed toward finding computationally effective
techniques to solve the integrated problem, and further simplification of the integrated
model.
References:


Chen, J., "Achieving maximum supply chain efficiency", IIE Solutions, June 1997


