MILK-RUN ROUTING AND SCHEDULING SUBJECT TO A TRADE-OFF BETWEEN VEHICLE FLEET SIZE AND STORAGE CAPACITY

Grzegorz Bocewicz, Wojciech Bożejko, Robert Wójcik, Zbigniew Banaszak

1 Koszalin University of Technology, Faculty of Electronics and Computer Science, Poland
2 Wrocław University of Science and Technology, Faculty of Electronics, Poland

Corresponding author:
Grzegorz Bocewicz
Koszalin University of Technology
Faculty of Electronics and Computer Science
Śniadeckich 2, 75-453 Koszalin, Poland
phone: (+48) 94 3478 717
e-mail: bocewicz@ie.tu.koszalin.pl

Received: 7 July 2019
Accepted: 22 August 2019

Abstract
The objective of the milk-run design problem considered in this paper is to minimize transportation and inventory costs by manipulating fleet size and the capacity of vehicles and storage areas. Just as in the case of an inventory routing problem, the goal is to find a periodic distribution policy with a plan on whom to serve, and how much to deliver by what fleet of tugger trains travelling regularly on which routes. This problem boils down to determining the trade-off between fleet size and storage capacity, i.e. the size of replenishment batches that can minimize fleet size and storage capacity. A solution obtained in the declarative model of the milk-run system under discussion allows to determine the routes for each tugger train and the associated delivery times. In this context, the main contribution of the present study is the identification of the relationship between takt time and the size of replenishment batches, which allows to determine the delivery time windows for milk-run delivery and, ultimately, the positioning of trade-off points. The results show that this relationship is non-linear.

Keywords
milk-run system design, routing, scheduling, trade-off point, Industry 4.0.

Introduction
The milk-run method is based on a transport route design under which a means of transport (a tugger train, a logistic train) is loaded and unloaded multiple times during one trip along a route that connects several workstations into a supply loop. This set-up allows a variable quantity of materials to be moved along a fixed route at regular time intervals, which means that the amount of materials can change but the delivery time always stays the same.

The main advantage of milk-runs, i.e. tours that are cyclically repeated according to a fixed schedule in a fixed sequence and with fixed arrival times to make frequent supplier-to-plant deliveries, is that they reduce variability by running on fixed schedules and thus allow to reduce inventory, capacity, and time buffers within a supply network. In other words, the use of the milk-run set-up allows a facility to increase the frequency of deliveries and synchronize transport and delivery of materials and products to the particular production stations, while at the same time ensuring that the number of means of transport used and the quantity of materials in stock are kept down to a minimum. In turn, however, requiring collision- and deadlock-free flow of tugger fleet implies the need to search for trade-off between fleet size and storage capacity. Therefore, the searching for a balance between the number of potentially conflicting tugger trains and a system’s storage capacity provides originality of the conducted research.
problem and motivation to solve it. Differently put, milk-run routing and scheduling problems boil down to determining in what time windows parts can be collected from suppliers and how many tugger trains and along which routes they should be transported to customers, so that the cost of transport operations and the size of the inventory in the supply chain are as low as possible. In the general case, however, the main goal is to simultaneously minimize fleet size and transportation and inventory costs [1, 2].

In this context, the search for an optimal periodic distribution policy, i.e., a plan on whom to serve, how much to deliver, and what regularly repeated routes to travel on by what fleet of vehicles, can be viewed as belonging to the class of Vehicle Routing Problem (VRP) [3, 4], which are NP-hard problems [5, 6]. Moreover, since the policy searched for must be such that a stock out is never caused at any of the customers, the maximum inventory level at the customers is never exceeded and the vehicle capacity is always satisfied, the problem considered is a capacitated and consistent VRP with multi-traffic pick-up and delivery with time windows and synchronization, which is a combination of variants of the VRP with multiple trips, the VRP with time windows, and the vehicles with capacitated constraints routing and delivery problem [7].

Milk-run set-ups are used in situations in which deliveries are made at regular intervals: in services (e.g. food delivery, waste collection, etc.) as well as in other sectors of economy, such as the automotive industry (e.g. vehicle assembly), home appliance industry (e.g. assembly of washing machines, fridges), etc. In the last of these application areas, it is perhaps most perspicious that the cyclic behavior of in-plant milk-run systems, i.e. periodic deliveries of materials (parts, subassemblies) [8], is dictated by the production takt time. The length of the takt time determines the size and the supply rhythm of replenishment batches of part-sets assembled at the assembly stations. In each takt time, assemblers at different workstations assemble parts from different part-sets. This means that the right part-sets should be delivered to workstations in batches large enough to guarantee an uninterrupted, rhythmical production flow. In general, the sizes of the part-set batches delivered to different workstations may be different. They depend, on the one hand, on the available buffer capacities, and, on the other, on the frequency of delivery of part-sets. In a special case in which part-sets are supplied one at a time, buffers need only have the capacity of one unit. Minimized buffer capacity requirements like this, however, come at the cost of the facility having to use numerous logistic trains to deliver supplies simultaneously to all workstations. Organized in this way, delivery is very expensive (as it requires the use of many logistic trains), and in the case of short production takt times, it is often impossible to complete (e.g. due to restrictions on train speed, topological limitations of transportation routes, traffic congestion, etc.).

The solutions encountered in practice concentrate either on establishing what minimum fleet of logistic trains can handle deliveries without exceeding the given buffer capacity, or on determining what buffer capacity is needed to store the replenishment batches delivered by the given fleet of tugger trains. In each of these cases, the routes and schedules for each train must, of course, be planned. Apart from this type of quantitative relationships connecting the parameters of a milk-run system structure (such as fleet size, number of carts pulled by each tugger train and their capacities, tugger train routes, etc.) there are also qualitative relationships between fleet size and storage capacity (understood as the total capacity of all workstation buffers). This means that the problem of delivery of materials to an assembly line with a given production takt time can be solved either by using more tugger trains, and hence a smaller storage capacity, or by using fewer tugger trains and a larger storage capacity.

It is easy to note that the replenishment batches delivered by a fleet of tugger trains and the batches of part-sets already stocked in buffers, together make up the work-in-progress (WIP) value. The minimum WIP value is determined by the trade-off points between fleet size and storage capacity. These trade-off points, in turn, depend on the adopted production takt time. A question that comes up naturally in this context is what the relationship is between the values of the production takt time and WIP. To solve this milk-run design problem, which is the main objective of this study, it is necessary to integrate problems of scheduling and dispatching of inventory in warehouses/supermarkets and production facilities with in-plant transport systems.

It is worth noting that research on the relationships between the constraints imposed by the structure of a production system and its performance capacity falls within the scope of Industry 4.0 [9–11], particularly with respect to the second of the four Design Principles of Industry 4.0 [11], i.e. information transparency, which stresses the need to seek solutions that allow information systems to create a virtual copy of the physical world, providing a framework for a context-aware information model.

Therefore, the paper contributes to the practice and methodology of management and produc-
tion engineering, as it enables fast online prototyping of supply schedules and transport routes of a tugger train fleet making adjustments for the trade-off points between fleet size and storage capacity. Since the use of commercially available software tools makes it possible to tackle practical-scale problems, hence it allows to create a virtual copy of the physical world, provides a programming framework for context-aware information model design while following digital tweens concept. Consequently such a model, when implemented in decision support system (DSS) software, supports correct responses to requests for decisions from the manufacturing process, especially regarding milk-run routing and scheduling.

The remaining part of the paper is structured as follows: Sec. 2 discusses related work on the models and methods for the milk-run design problem; Sec. 3 presents an illustrative example, describes a declarative model of milk-run design and gives a formulation of the problem. The results of numerical experiments are discussed in Sec. 4 and the main points of the study are summarized in Sec. 5.

Literature review

In-plant milk-run systems are transportation systems in which materials are delivered from a central storage area to several points of use on fixed routes at short, defined intervals [12]. The benefits of using a system of this type include improved efficiency of the overall logistics system and potential substantial savings in shortening the total distance travelled and minimizing the number of vehicles used, along with considerable financial benefits related to reduced inventory costs [13]. Many case studies discuss the structure and operation of milk-run systems used in various industries such as the automotive industry [14], furniture industry [8], food delivery [15], and waste collection [16]. The literature related to milk-run design problem is extremely diverse: it ranges from optimal lot sizing, through network design and vehicle routing to publications specifically addressing lean manufacturing. The most commonly formulated routing problems are those aimed at maximizing the utilization of fleet capacity, finding the best routing and determining the number of parts to be collected from each supplier on each trip. Other frequently encountered routing problems address the questions of "How to assign certain sequences of stops to certain routes?" and "How to configure tugger trains?" [17]. Only a limited number of papers are devoted to robust and congestion-free scheduling of a fleet of vehicles subject to in-plant layout constraints. In this respect, the most relevant factors are those which depend on critical, and often unpredictable, traffic congestions which occur when logistics operators allocate too many collecting tasks to the available vehicles, generating unperformed activities due to assumed just-in-time constraints imposed by the time windows of customer services [18, 19]. Methods that are most commonly employed for such purposes include those that use the formalism of max-plus algebra [20, 21] and constraint programming [22], as well as computer simulation [12]. The main research gap with respect to the milk-run design problem is the lack of an explicit, deterministic model representing inventory behavior in the case of cyclically repeated delivery profile schedules with levelled lot sizes. Furthermore, the literature offers no operational model that would simultaneously take into account transport concept assignment, frequency assignment, and the milk-run scheduling decision [13, 17, 23, 24]. Milk-run problems of component/part/commodity distribution can be classified similarly to the exhaustively studied extensions of the VRP, such as the capacitated and consistent VRP with multi-trip multi-traffic pick-up and delivery problem with time windows and synchronization which is a combination of variants of the VRP with multiple trips, VRP with time windows, and capacitated routing and delivery problem [7, 13, 22, 23, 25, 26]. The modeling frameworks used for the modelling and analysis of milk-run set-ups include operation research methods (such as linear and non-linear programming, MLP, computer simulation, and so on) and artificial intelligence methods, such as evolutionary computation (including metaheuristic and stochastic optimization algorithms) [19, 23].

Given this background, our contribution comes down to assessing the possibility of using the declarative modeling framework for in-plant milk-run traffic systems. The main objective of our study is to analyze the relationship linking the trade-off between fleet size and storage capacity with implied milk-run routes and schedules. This issue has been partially discussed in our previous work on the reference model of a milk-run delivery problem [7] and a declarative modeling framework for congestion-free routing and scheduling of tugger trains [27].

Illustrative example

The assembly line under consideration, whose route passes through 60 stations \( S = \{ SN_1, ..., SN_{60} \} \) is shown in Fig. 1. For simplicity’s sake, it is assumed that all types of products assembled on this line \( W_j, j \in \{1, ..., k \} \) are produced along the same technological route and at the same takt time \( T = 130 \) [s].
The elements of each batch of products are moved between the neighboring stations by a gantry robot. The product is assembled in steps at different assembly stations. During the assembly, part-sets are assembled simultaneously, at the same takt time, on all stations of the line. The part-sets are delivered to the stations in containers of

\[ PS = \{PS_{a1}^{a1}, ..., PS_{a1}^{a60}, ..., PS_{a60}^{a60}\} \]

where: \( PS_{ai}^{aj} = \{c_{p}, ..., c_{a}\} \) – a part-set of elements necessary to assemble the \( a_i \in \{1, ..., k\} \)-th product at assembly station \( SN_i \).

The containers are delivered to stations by three logistic trains \( F = \{F_1, F_2, F_3\} \) pulling two carts each with a capacity of 150 part-sets. Assembly sta-
tions are supplied with parts in batches, in quantities that are multiples of part-set supplies (where \( d \) is the batch size). For example, if the transport route of a given logistic train passes through two stations, \( SN_i \) and \( SN_j \), which assemble parts from part-sets \( PS_i = \{ c_p, ..., c_a \} \) and \( PS_j = \{ c_g, ..., c_v \} \), respectively, then this train delivers to assembly station \( SN_i \) container \( BN_i \) which carries a batch of \( d \) elements of part-set \( PS_i \) and container \( BN_j \) with a batch of \( d \) elements of part-set \( PS_j \). In other words, container \( BN_i \) contains \( d \cdot c_p, ..., d \cdot c_a \) elements \( \{ c_p, ..., c_a \} \) and container \( BN_j \) contains \( d \cdot c_g, ..., d \cdot c_v \) elements \( \{ c_g, ..., c_v \} \). For the sake of simplicity, it is assumed that the numbers, volume and weight of the elements are the same for all containers.

Assembly stands are equipped with buffers \( B_l \) (\( B_1 \) – warehouse, \( B_2 \)–\( B_{10} \) workstation buffers). Buffers can be shared by several stations; \( r_k \) is the number of stations using the same buffer \( B_k \). The capacities of these buffers are multiples of the part-sets delivered to their respective stations. If two assembly stations, \( SN_i \) and \( SN_j \), which assemble parts from containers \( BN_i \) and \( BN_j \), respectively, use the same buffer \( B_l \), then the capacity of that buffer corresponds to the total number of part-sets delivered to these stations. In the case under consideration, capacity \( Z_l \) of buffer \( B_l \) is \( Z_l = 2 \cdot d \), because each container delivered to a station contains a batch of part-sets comprising \( d \) elements. A transportation system organized in this way works under the assumption that containers are delivered to buffers by trains and then transferred to stations by hand or by gravity feeders. Beside each buffer, there is a special bay where trains can stop to unload the containers in a safe, collision-free manner. It is assumed that the durations of tugger train trips are known (Table 1) and also known is function (1) which describes the unloading times \( tc_k^l \) of part-sets (unloading time of \( k \)-th train at \( l \)-th buffer):

\[
    tc_k^l = \begin{cases} 
        60 & \text{for } l = 1, \\
        \epsilon_k^l & \text{for } l > 1,
    \end{cases} 
\]

where \( \epsilon_k^l \) – number of part-sets delivered to buffer \( B_l \) by train \( F_k \).

The routes of tugger trains \( \{ \pi^1, \pi^2, \pi^3 \} \) are marked by sequences of the buffers served by these trains. For example, in accordance with Fig. 1, train \( F_1 \) delivers containers along route \( \pi^1 = (B_1, B_9, B_{10}, B_4) \), train \( F_2 \) travels along route \( \pi^2 = (B_1, B_8, B_5, B_6, B_7) \), and train \( F_3 \) carries containers along route \( \pi^3 = (B_1, B_5, B_3, B_2) \).

The routes allow cyclic replenishment of buffer stocks over a period of \( TT = 650 \) [s] – see the schedule in Fig. 2. Replenishment batches \( d = 5 \) delivered to each buffer last for 5 takt times. This means that the total capacity of the buffers used, i.e. the system’s storage capacity, cannot be less than 300 part-sets \( \{ |B_2| = 50, |B_3| = 25, |B_4| = 25, |B_5| = 50, |B_6| = 25, |B_7| = 25, |B_8| = 25, |B_9| = 25, |B_{10}| = 50 \} \).

In the general case, solutions for this system may feature different numbers of trains \( K = |F| \) and different storage capacities \( SB \), and also (as a consequence) different sizes of replenishment batches \( d \). It is easy to see that the values of these parameters are dependent on one another. For example, if fewer trains are used, then delivery periods \( TP \) become longer, making it necessary to increase buffer capacity \( SB \). And conversely, a reduction of the admissible capacity of the buffers \( SB \) necessitates the use of a larger number of tugger trains \( K \).

In general, when the travel times of logistic trains and the production takt time \( T \) are known, one can determine trade-off points which simultaneously minimize fleet size \( K \) and storage capacity \( SB \). In other words, what is sought are solutions for delivery plans and the associated pairs \( CS = (K, SB) \) in which the values of fleet size and storage capacity are chosen in such a way that a reduction in the values of one of them results in an increase in the value of the other and vice versa. Parameters \( K \), \( SB \) depend on \( d \). These relationships for the set-up shown in Fig. 1 are illustrated by the graphs in Fig. 3. It can be seen, for example, that when batch size is chosen to be \( d = 6 \), an admissible delivery plan requires the use of two logistic trains \( (K = 2) \) in a system with a total buffer capacity of \( SB = 360 \) part-sets.

| Table 1: Travel times of tugger trains along sectors \( S_i \) |
|-----------------|---|---|---|---|---|---|---|
|                | \( S_1 \) | \( S_2 \) | \( S_3 \) | \( S_4 \) | \( S_5 \) | \( S_6 \) | \( S_7 \) |
| Travel times \([s]\) | 30 | 40 | 20 | 30 | 20 | 40 | 50 |
| Travel times \([s]\) | 40 | 30 | 40 | 30 | 40 | 40 | 40 |

Management and Production Engineering Review

Volume 10 ● Number 3 ● September 2019

45
Fig. 2. Schedule for delivery of replenishment batches to a system with a storage capacity of $SB = 300$ part-sets. The part-sets are delivered in batches of $d = 5$ by a fleet of $K = 3$ tugger trains.

Fig. 3. Relationship between trade-off points (specified by the pairings: fleet size $K$ and storage capacity $SB$) and batch size $d$. 
In the example shown, there are three trade-off points: pair CS1: K1 = 5, SB1 = 240, pair CS2: K2 = 3, SB2 = 300, and pair CS3: K3 = 2, SB3 = 360, which are returned for delivery batches with respective sizes of d1 = 4, d2 = 5, and d3 = 6.

It is easy to see that the choice of any other solution requires either the use of a larger number of trains (K /) or increasing the total storage capacity of the buffers (SB /). A delivery schedule corresponding to solution CS2 is shown in Fig. 2.

The trade-off points correspond to alternative delivery set-ups in which different sizes of the replenishment batches minimize WIP. In the case under consideration, increases in fleet size K are compensated by decreases in storage capacity SB and vice versa. Changes in the size of replenishment batches in the range d = (6, 5, ..., 10) which can be delivered by a fleets of K = (2, 3, ..., 6) trains result in a respective increase in storage capacity in the range SB = (400, 460, ..., 600). It is worth noting that in the next range of changes in the size of replenishment batches d = (11, 12, ..., 15) which can be delivered by fleets of K = (3, 4, 5, 6) trains, the changes result in a further increase in the storage capacity value in the range SB = (660, 720, ..., 900). An analogous situation takes place in the next range of changes in the size of replenishment batches d = (16, ..., 20), which can be delivered by fleets of K = (4, 5, 6) trains: the changes result in a further increase in storage capacity in the range from SB = [960, 1020, ..., 1200]. The cases of increases in fleet size K described above are explained by a constant which is constrained by the capacity of the trains. By adding additional carts to the trains, i.e. increasing their capacity, one can keep the number of trains constant.

To sum up, when the travel times of logistic trains and the storage capacities of buffers and trains are known, the value of production takt time allows to determine both the minimum and maximum number of tugger trains needed. Knowing this range, which defines how many trains can be used at the same time, one can determine the respective admissible range of total buffer capacities. A question that naturally comes to mind in this context is the question of the relationship between takt time T and the corresponding trade-off points, i.e. pairs of minimum size of a tugger train fleet K and minimum storage capacity SB.

### Parameters

- **G** graph of a transportation network \( G = (N, E) \); \( N = \{1, \ldots, n\} \) is a set of nodes, \( E = \{(i, j)\} \) is a set of edges \( B = \{B_1, \ldots, B_{SB}\} \subseteq N \) - nodes representing buffers; \( S = \{S_1, \ldots, S_{SB}\} \subseteq N \) - nodes representing sectors, \( B \cup S = N \), \( B \cap S = \emptyset \).
- \( z_i \) demand for each node (for the case when \( d = 1 \)): \( z_{B_1} \in N \), \( z_{B_3} = 0 \), \( z_{S_1} = 0 \).
- \( t_i \) time spent (pick-up/drop-off or transport) on node \( i \).
- \( F \) set of transportation tugger trains, \( |F| = K \).
- \( FL \) maximum loading capacity of each tugger train.
- \( TS \) minimum time between successive train mission start times.
- \( H \) time horizon \( H = [0, TT] \).
- \( T \) production takt time.

### Decision Variables

- \( TT \) deadline, size of the time window.
- \( d \) batch size (the notation for batch size is shorthand for a more specific notation: \( d_i \) which defines the batch size of part-sets of the \( j \)-th product delivered to the \( i \)-th station).
- \( SB \) storage capacity.
- \( Z_i \) capacity of node \( B_1 \).
- \( K \) fleet size \( K = |F| \).
- \( x_{i,j}^k \) binary variable used to indicate if the \( k \)-th tugger train travels from node \( i \) to node \( j \).

\[
x_{i,j}^k = \begin{cases} 1 & \text{if } k\text{-th tugger train travels from node } i \text{ to node } j \\ 0 & \text{otherwise} \end{cases}
\]

- \( s \) start time of the \( k \)-th tugger train: \( s \in N \).
- \( y_i^k \) time at which the \( k \)-th tugger train arrives at node \( i \): \( y_i^k \in N \).
- \( c_i^k \) payload weight delivered to node \( i \) by the \( k \)-th tugger train.
- \( t_e_i^k \) real time spent (pick-up/drop-off or transport) on node \( i \) by the \( k \)-th tugger train.
- \( \pi^k \) route of the \( k \)-th tugger train

\[
\pi^k = \{(v_1, \ldots, v_i, v_{i+1}, \ldots, v_n), v_i \in N, x_{v_i,v_{i+1}}^k = 1\}.
\]
Sets

- Y^k: set of times \( y_{i}^{k} \) – schedule of the k-th tugger train,
- \( Y\): family of \( Y^k \) – schedule of tugger train fleet,
- \( C^{k} \): set of \( c_{i}^{k} \) – payload weight delivered by the k-th tugger train,
- \( C \): family of \( C^{k} \),
- \( \Pi \): set of UAV routes \( \pi^{k} \),
- \( SM \): mission \( SM = (\Pi, Y, C) \).

Constraints

**Arrival time at nodes.** Relationship between the binary decision variables \( x_{i,j}^{k} \) and variables \( y_{i}^{k} \) (s^k):

\[
(x_{i,j}^{k} = 1) \implies (y_{j}^{k} = s_{i}^{k} + t_{i}^{k}), \quad \forall (i, j) \in E, \quad k = 1 \ldots K, \quad (x_{i,j}^{k} = 1) \implies (y_{j}^{k} = y_{i}^{k} + t_{i}^{k}), \quad \forall (i, j) \in E, \quad k = 1 \ldots K,
\]

\[
y_{i}^{k} + t_{i}^{k} \leq TT \sum_{j \in N} x_{i,j}^{k}, \quad k = 1 \ldots K,
\]

\[
|s_{i}^{k} - s_{v}^{k}| \geq TS, \quad k, v = 1 \ldots K, \quad k \neq v,
\]

\[
TT \leq d \times T, \quad \text{Mod}(TT, T) = 0.
\]

**Collision avoidance.** Shared nodes cannot be used at the same time when they are occupied by two tugger trains (\( x_{i,j}^{k} = 1 \) and \( x_{\alpha,j}^{k} = 1 \)):

\[
\left[ (x_{i,j}^{k} = 1) \land (x_{\alpha,j}^{k} = 1) \right] \implies \left[ (y_{j}^{k} + t_{i}^{k} \leq y_{\alpha}^{k}) \lor (y_{\alpha}^{k} + t_{i}^{k} \leq y_{j}^{k}) \right],
\]

\[
k, v = 1 \ldots K, \quad k \neq v, \quad (i, j), (\alpha, j) \in E.
\]

**Capacity.** The demand assigned to a tugger train should not exceed its capacity

\[
c_{i}^{k} \geq 0, \quad \forall i \in N,
\]

\[
c_{i}^{k} \leq FL \times \sum_{i \in N} x_{i,j}^{k}, \quad \forall j \in N, \quad k = 1 \ldots K,
\]

\[
\sum_{i \in N} c_{i}^{k} \leq FL, \quad k = 1 \ldots K,
\]

\[
Z_{i} = d \times z_{i}, \quad \forall i \in N,
\]

\[
\sum_{k=1}^{K} c_{i}^{k} = Z_{i}, \quad \forall i \in N,
\]

\[
t_{i}^{k} = t_{i} + c_{i}^{k}, \quad \forall i \in B,
\]

\[
t_{i}^{k} = t_{i}, \quad \forall i \in S,
\]

The sum of all weights \( c_{j}^{k} \) carried by tugger trains should not exceed the maximum carrying payload \( FL \) and demand at node \( d \cdot z_{i} \).

**Flow of tugger trains.** A train that arrives at a node must leave from that node

\[
\sum_{j \in N} x_{i,j}^{k} = \sum_{j \in N} x_{i,j}^{k}, \quad \forall i \in N, \quad k = 1 \ldots K,
\]

\[
\sum_{j \in N} x_{i,j}^{k} \leq 1, \quad \forall i \in N, \quad k = 1 \ldots K,
\]

\[
\sum_{k=1}^{K} \sum_{i \in N} x_{i,j}^{k} \geq 1, \quad \forall i \in B.
\]

**Cost function.** The cost function is defined as the pair of fleet size \( K \) and storage capacity \( SB \):

\[
CS = (K, SB).
\]

**Problem formulation**

To solve the problem under discussion, one has to answer the following question:

Consider a transportation network \( G \) representing a production system with a production takt \( T \). Does there exist mission \( SM \) (determined by variables \( \Pi, Y, C \)) which can minimize \( CS \) (minimize fleet size \( K \), and storage capacity \( SB \)) under constraints (2)–(20)?

The investigated problem can be viewed as a Constraint Optimization Problem (COP) given by (22):

\[
\text{COP} = (V, D, C, CS), \quad (22)
\]

where \( V = \{ \Pi, Y, C \} \) – a set of decision variables including: \( \Pi \) – a set of train routes, \( Y \) – a schedule of a tugger train fleet, \( C \) – a set of payload weights delivered by the tugger trains, \( D \) – a finite set of decision variable domain descriptions, \( C \) – a set of constraints specifying the relationships between tugger train routes, tugger train schedules and transported materials (2)–(20), \( CS \) – a cost function representing the pair: fleet size \( K \), buffer capacities \( SB \) (22).

To solve COP, one must determine such values of the decision variables from the adopted set of domains for which the given constraints are satisfied and the cost function reaches its maximum. By implementing a COP in a constraint programming environment such as IBM CPLEX, one can construct a computational engine that can be implemented in an interactive decision support system (DSS).
Computational experiments

Consider the transportation network from Fig. 4a, which follows the milk-run layout from Fig. 1. In the example considered earlier, this system operated at a production takt time of $T = 130$ s. The trade-off determined for this takt time (see $CS_2 = (3, 300)$ from Fig. 3) assumes that goods in this system are transported by a fleet of $K_2 = 3$ trains (each train pulling two carts with a capacity of 150 part-sets, see Fig. 4b) and the total buffer capacity of the system is $SB = 300$ part-sets. The delivery schedule is shown in Fig. 2.

The goal of the experiments was to determine the trade-off points for different values of production takt time $T$, in particular those in the range of 70–200 [s]. In other words, we looked for an answer to the question below for selected $T$ values: *Does there exist mission $SM$ (determined by variables $\Pi, Y, C$) which can minimize $CS$ (minimize fleet size $K$, and storage capacity $SB$) for the given production takt time $T$?*

To find the answer, COP (21) was solved for $T = (70, 90, \ldots, 200)$. The problem was implemented and solved in the constraint programming environment IBM ILOG CPLEX (Windows 10, Intel Core Duo2 3.00 GHz, 4 GB RAM). The results of the experiments are summarized in Table 2. To make the presentation of the results more legible, the trade-off points are listed in rows corresponding to the same fleet size.

![Diagram](image_url)

**Legend:**
- two-way route
- one-way route

Fig. 4. Diagram showing (a) a transport network for the system of Fig. 1 and (b) transfer of part-sets from train $F_1$ to stations $SN_{46} - SN_{60}$.

<table>
<thead>
<tr>
<th>Takt time $T$</th>
<th>70</th>
<th>90</th>
<th>110</th>
<th>130</th>
<th>150</th>
<th>170</th>
<th>200</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trade-off $CS = (K, SB)$ $K$</td>
<td>2</td>
<td>(2,720)</td>
<td>(2,480)</td>
<td>(2,360)</td>
<td>(2,360)</td>
<td>(2,240)</td>
<td>(2,240)</td>
</tr>
<tr>
<td>3</td>
<td>(3,540)</td>
<td>(3,420)</td>
<td>(3,300)</td>
<td>(3,300)</td>
<td>(3,300)</td>
<td>(3,240)</td>
<td>(3,180)</td>
</tr>
<tr>
<td>5</td>
<td>(5,480)</td>
<td></td>
<td>(5,240)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>(6,360)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The results are shown in graphical form in Figs 5 and 6. The decrease in storage capacity $SB$ – visible in Fig. 5 – caused by an increase in production takt time $T$ is a consequence of an extension of the time window $TT$ in which replenishment batches are delivered. The same increase in takt time $T$ leads, for the same reason, to a reduction in maximum fleet size $K$.

It is worth noting that the different trade-off points $CS = (K, SB)$ for the same takt time are achieved by using different tugger train routings. For example, the alternative trade-off points $CS' = (2, 360)$ and $CS'' = (3, 300)$ implicated by takt time $T = 110 \text{ [s]}$, are returned by different tugger train routings, as illustrated in Figs 7a and 7b.
The different routings are associated with distinct delivery costs, which is a consequence of differences in the total distance travelled, etc.

The routes are planned taking into account the constraints adopted in the model and assuming that the resource conflicts which lead to deadlocking of processes are resolved using a deadlock prevention method, thus implementing conditions which guarantee avoidance of congestion [7, 27]. This means that the obtained set of trade-off points can be searched for solutions that would minimize this type of costs. This is particularly important in cases when trains travel and return multiple times along the same route sections.

The experiments confirm that there exist relationships between takt time and the size of replenishment batches, which, in turn, determine the time windows for milk-run delivery, and, ultimately, the positioning of trade-off points. Among others, the results indicate that there is a non-linear relationship between takt time $T$ and the values characterizing the trade-off points $(K, SB)$. In particular, they demonstrate that a reduction in the production takt time must be accompanied by an increase in fleet size $K$ and storage capacity $SB$.

The data we obtained also point to a connection between the parameters of the structure of a production system (such as route structure, fleet size, load capacity of tugger carts, the number, distribution and capacity of buffers, etc.) and the performance characteristics of the system implicated by this structure (such as batch size, tugger train routes, delivery cycles of the individual tugger trains, etc.). The existence of such a connection means that one can search for desired system behaviors (e.g. the efficiency of the milk-run system being designed) given the limitations of the available system structure, or, the other way round – one can search for structural parameters (e.g. fleet size) which will ensure the desired operation of the system. This means that it is possible to build a decision support system for the design of milk-run systems in which one could alternately look for answers to the questions below in an interactive mode [28]: Does there exist a solution that would allow replenishment deliveries to be made at the expected efficiency level? What parameters and/or characteristics of system structure elements should be changed to ensure that replenishment batches are delivered at the expected efficiency level?

**Conclusions**

The novelty of this study is that it proposes an integrated modelling approach to milk-run system design and operation which takes into account the relationships linking the takt time of production
flow with the size of the replenishment batches and the positioning of trade-off points. The declarative framework that stands behind the reference model of the milk-run design problem formulated as a CSP model is a useful tool which allows to handle those relationships. Implemented computationally, it enables fast online prototyping of supply schedules and transport routes of a tugger train fleet making adjustments for the trade-off points between fleet size and storage capacity. The use of commercially available software tools, such as CPLEX/ECLPS/Gurobi, etc., which make it possible to tackle practical-scale problems, can be viewed as an attractive solution for problem-oriented DSS. This means that our study, being in line with the concept of Industry 4.0, which stresses the need to seek solutions that allow information systems to create a virtual copy of the physical world, provides a programming framework for context-aware information model design. The results of the tests demonstrate that the proposed reference CSP model is a useful tool that allows one to formulate problems of milk-run routing and scheduling subject to a trade-off between vehicle fleet size and storage capacity.

Aside from the research perspective presented in this paper, other directions of study worth mentioning are those aimed at investigating the conditions that would allow planners to reschedule milk-run flows according to customers’ changeable demands, i.e. production flows. Other potentially interesting areas of future investigation relate to the smooth transition between two successive cyclic steady states corresponding to the current and rescheduled flows.

The paper was partially supported by the National Science Centre Poland, grant OPUS no. DEC 2017/25/B/ST7/02181.

References


