Case Study

Network planning for scheduling operations in air cargo handling: a tool in medium term goodsflow control

M.A.A.P. Verwijmeren and C.B. Tilanus

Graduate School of Industrial Engineering and Management Science, Eindhoven University of Technology, Postbox 513, 5600 MB Eindhoven, Netherlands

Received August 1992; revised December 1992

Abstract: In the hub of a hub-and-spokes network for airfreight transportation, the main part of the incoming and outgoing goodsflow is in special loading units for airfreight. These loading units are metal pallets and containers up to eighteen cubic meters in size. The key part of operations in the hub is breaking down incoming loading units and building up outgoing loading units. These operations are subject to resource restrictions (limited number of platforms and manpower teams) and time restrictions (between arrival times of flights and departure times of flights). With some substitutions, standard software for network planning is successfully applied to scheduling transshipment operations of loading units.

Keywords: Airfreight; Network planning; Scheduling; Standard software; Goodsflow control

1. Environment

The cargo division of an international airline company, KLM Royal Dutch Airlines (KLM), is responsible for logistic services related to airfreight. These services include fast and reliable transport of air cargo as well as timely handling of air cargo on the ground.

Schiphol Airport near Amsterdam is the homebase of KLM and the centre of its global network. The centre ('hub') has connections ('spokes') with stations all over the world. The structure of the hub-and-spokes network is depicted in Figure 1.

The cargo division is dependent on the two other divisions of KLM. The passenger division is responsible for passenger transportation and emphasizes passenger customer service. The operations division is responsible for physical transportation operations of both passengers and cargo and confronts the cargo division and the passenger division with the capacity limitations in transportation by air. All aircraft are used for 'combined' carriage, transporting passengers (with their baggage) as well as air cargo, in variable proportions. Thus, two divisions are competing for common limited flight capacity. Conflicting interests are taken into account twice a year when the so-called timetable is determined. The timetable states all flights of KLM for the next half year. So within a framework of about six months planned arrivals and departures of aircraft are completely fixed. While drafting the
timetable, the flight capacity is split into passenger capacity and cargo capacity.

For continental (within Europe) air cargo transportation not only aircraft but also trucks are used. ‘Flights’ carried out by trucks are more flexible than flights by aircraft. If capacity is redundant, trucks can be cancelled. If capacity runs short, extra trucks can be arranged.

2. Goodsflow

The goodsflow in airfreight can be divided in two parts. The first part is called ‘transportation part’. In this part logistic units are transported by air or by road. The logistic units in the transportation part are aircraft or trucks and ULD’s. ULD stands for ‘Unit Load Device’ and is the main loading unit in airfreight. ULD’s can be either aluminium pallets or aluminium containers, having a load volume of maximum eighteen cubic meters. The loaded weight in a ULD is rarely constraining.

The second part is called ‘handling part’. In that part logistic units are handled on the ground. The logistic units in the handling part are ULD’s, shipments and packages, depending on the processes in the handling part. ULD and package are physical logistic units; shipment is an administrative logistic unit. A shipment consists of one or more packages with the same origin and destination specification. Packages of one shipment are transported on the same flight, but may be loaded on different ULD’s. When a shipment enters the network all flights from origin to destination are determined in accordance with the client’s wishes.

Between arrival of a ULD in the hub and departure of a ULD from the hub two main processes take place. First, incoming ULD’s are broken down. In the breakdown process ULD’s are ungrouped into packages. Later, outgoing ULD’s are built up. In the buildup process packages are grouped into ULD’s. However not all packages and ULD’s pass through these processes.
At a station of origin packages are grouped into ULD's. The ULD's are transported to the hub. In the hub some incoming ULD's are transhipped onto outgoing flights without being broken down. The other incoming ULD's have to be broken down. The percentage of ULD's of a flight that has to be broken down is called breakdown fraction.

In Figure 2 a frequency distribution for the breakdown fraction is shown. There is a large portion of flights with an extremely low or an extremely high breakdown fraction (19% of the flights have a breakdown fraction below 10%; 16% of the flights have a breakdown fraction of over 90%). These varying breakdown fractions have their impact on the demand for handling capacity in the hub and the stations. (The overall average breakdown fraction over all flights is about 40%).

Packages on a ULD that has to be broken down are either directly transferred to consignees (leaving the network) or prepared for different flights from the hub to stations of final destination (proceeding in the network). The first group of arriving packages only passes the breakdown process. The second group passes both the breakdown process and the buildup process. Packages that are delivered into the hub by shippers (entering the network) only pass the buildup process.

3. Goodsflow control

A concept of goodsflow control in airfreight is shown in Figure 3. In the concept the term of goodsflow control is related to the distinguished logistic units. The logistic units discerned in the goodsflow control are respectively aircraft and truck for the long term, ULD for the medium term and shipment and package for the short term.

The long term is over six months, because that is beyond the scope of the flight timetable. The short term is less than two days, because only then shipment and package data are available. In the medium term (between two days and six months), flights are fixed, but load data like breakdown fraction of ULD's and deadlines are based on forecasts.

4. Problem definition

Goodsflow control can be classified with six cells, by discerning on the one hand the part in the goodsflow (transportation or handling) and on the other hand the term of the goodsflow control (long, medium or short). From that point of view this study falls under the medium term goodsflow control in the handling part.

Both for the long term and the short term goodsflow control systems are available in KLM. So far the medium term goodsflow control, having a term from two days to six months, has hardly been paid attention to. The medium term goodsflow control however is important, because operations planning and manpower planning decisions are made for the associated time window.

One group of decisions concerns the availability of resources. How should the permanently available manpower be scheduled? Should temporary manpower be hired? Another group of decisions is related to the demand for resources. What is the impact of airfreight characteristics (e.g., the breakdown fraction) on the demand for resources? How should airfreight characteristics be in order to prevent insufficiency of resources in processes?

The intention of this study is to systematically improve a part of the goodsflow control; for a general discussion of operational planning in airline business with the help of operations research, see [1]. In this study, a model is built for the medium term goodsflow control in the handling part. The central problem in the study is choosing an appropriate way of scheduling opera-
tions in the breakdown process and the buildup process.

For the medium term goodsflow control the smallest logistic unit that is distinguished has to be determined. In other words: the aggregation level of the system that is the basis for the model has to be specified.

Goodsflow is controlled on the aggregation level of the ULD's. On the one hand, using aircraft and trucks as units for handling capacity norms is unacceptable because the results would be too rough. On the other hand, using shipments and packages as units for handling capacity norms is judged unnecessarily detailed and practically infeasible. In the past, there has been an unsuccessful attempt at KLM to schedule operations on the aggregation level of shipments and packages. That attempt failed because both input and computing time became prohibitive. By working on the level of the ULD's, the interrelations between ULD's, shipments and packages can be neglected, which reduces the model complexity drastically.

Both in the breakdown process and in the buildup process operations on ULD's are carried out. An operation in the breakdown process is defined as the breakdown of a ULD, an operation in the buildup process as the buildup of a ULD. The control model is divided into a model for simulating breakdown or buildup operations and a model for scheduling breakdown or buildup operations. These models are discussed in the next section.

5. Simulating and scheduling operations

The operations that have to be carried out in the processes are simulated. A top-down approach is chosen for the simulation. ULD's are definitely not simulated by first simulating shipments (packages and their interrelations) and then simulating the way these shipments are grouped into ULD's. Because of the large amount of data to be processed this bottom-up approach is eliminated.

Simulation of ULD's takes place by combining the simulation of aircraft and truck movements ('flights') with the simulation of so-called ULD parameters. ULD parameters are indices of the ULD goodsflow, for example number and kind of ULD's expected in a flight, expected fraction of ULD's that have to be broken down or built up, durations of breakdown or buildup operations and time restrictions for these operations. So data about future flights are combined with data about ULD's.

The simulated series of operations has to be scheduled. The scheduling of operations is subject to two kinds of constraints: time restrictions and resource restrictions. In this case time restrictions are clock time restrictions: not relative time intervals but absolute points of time are stated.

For each operation two time restrictions apply: an earliest start and a latest finish.

In the breakdown process the earliest start of an operation depends on the arrival time of a flight and the time that elapses between arrival and availability in the breakdown process. The latest finish, or deadline, of an operation in the breakdown process depends on the moment that one of the packages in the ULD is needed in other processes, for example in the buildup process.

In the buildup process the two time restrictions are determined analogously. The earliest start depends on the availability of packages in the buildup process, the latest finish is directly related to departure times of flights.

The scheduling of operations is not only subject to time restrictions but also to resource restrictions. Resources in the processes have to do with manpower and location. Manpower resources consist of breakdown teams and buildup teams. These are teams of four persons together responsible for respectively ungrouping a ULD into packages and grouping packages into a ULD. Location resources are breakdown platforms and buildup platforms. A platform is a rectangular metal surface having a variable vertical position on which a ULD is broken down or built up. To accomplish an operation one team and one platform are required.

The situation in which operations, having both time restrictions and resource restrictions, have to be scheduled, can be interpreted in two different ways.

The situation can be interpreted as a job shop situation in which a team and a platform together form a workstation and in which jobs are identical to the simulated operations. Such a job shop
situation with resource restrictions and time restrictions can be translated in a queuing model with a limited number of servers. For scheduling the jobs (operations) decision rules dealing with restrictions and priorities have to be developed. Unfortunately, there exists no easily accessible standard software for the job shop situation [2].

In this case study, the situation is interpreted as a project instead of a job shop. Then operations are considered as activities in the project. The activities (operations) are subject to time restrictions and resource restrictions. The start of the project is some given clock time and the finish of the project is another clock time; between these two clock times given operations have to be carried out.

The project is modelled in a network in which activities are not directly interrelated, that means there are no precedence relations between activities in the network. Figure 4 gives the representation of the network. The operations that are represented in a network can be scheduled with the help of network planning with clock time and resource restrictions.

Standard software for network planning is abundant. In this study the network planning is implemented in SAS (Statistical Analysis System) [4], which is the KLM standard software used. SAS is equipped with a module for operations research. Network planning in SAS uses the ‘Critical Path Method’ (CPM).

In SAS only one clock time restriction per activity can be imposed, instead of the required two time restrictions. Therefore in this case study each activity is split into a core activity and a dummy activity. The dummy activity is successor of the core activity and takes care of the second time restriction.

The simulation of operations described earlier also takes place in SAS. Operations having time restrictions are simulated. Per shift per working day the available number of teams and platforms are simulated as well. In other words the complete network, including resource restrictions, is simulated.

6. Active model

Simulation and network planning are integrated in one automated model. ‘Active model’ stands for the complete model in SAS fed with real life data. Some of the aspects of the active model are discussed in this section.

The ‘control period’ is the period for which the model controls the goodsflow. The control period has a length of seven days and is extended by a ‘phase in period’ (preceding the control period) and a ‘phase out period’ (succeeding the control period), both having a duration of three days. So a system in balance is approximated. Phase in and phase out periods of both seven days were initially intended, but turned out to be too computer capacity consuming in the given computer environment. Experience so far has shown that three days for phase in and phase out are adequate.

For the total period (phase in period, control period and phase out period) about 2350 operations are simulated (approximately 1250 operations per week). Due to splitting activities the network consists of about 4700 activities. After choosing the control period, the model automatically simulates all operations and generates the complete network.

Network planning under restrictions can be either resource constrained or time constrained. In resource constrained scheduling, resource restrictions have priority over time restrictions. In time constrained scheduling, time restrictions have priority over resource restrictions. If resource shortages occur during time constrained scheduling, extra, fictitious resources are employed.

In this study time constrained scheduling is chosen, so operations are not allowed to be delayed. If operations were delayed, the utilization
Network planning in SAS uses the so-called parallel method to schedule activities. First activities are scheduled without taking resource restrictions into account (a forward pass followed by a backward pass). For each activity its earliest start and latest finish are calculated. Next, it is attempted to schedule activities as early as possible within their float, in order of earliest start, taking resource restrictions into account.

A time pointer gets a value equal to the overall earliest start of all activities. Activities that are candidates for scheduling, i.e. have an earliest start equal to the time pointer, are selected and ranged in order of ascending remaining float.

If sufficient resources are available or the float is not positive, the actual candidate is scheduled immediately. If the availability is insufficient and there is still positive float, the actual candidate is not scheduled but deferred (later on such an activity will become a candidate for scheduling again). Then the next candidate is considered.

If all candidates have been examined, for all activities that have not been scheduled yet, the earliest start and latest finish are recalculated. The new earliest start for activities that have already been candidates is set equal to the time at which the next change in number of free resources occurs.

Then new candidates are determined with the help of the changed value of the time pointer that is equal to the new overall earliest start of the remaining activities. In this way activities are scheduled as early as possible.

In other projects activities are normally scheduled as early as possible, within their float and satisfying resource restrictions. However, in this case there are good reasons for scheduling activities as late as possible under the same conditions.

For several reasons KLM schedules as late as possible. By scheduling buildup operations just before flight departures (as late as possible) ULD's can be built up more easily. If buildup operations are scheduled as late as possible, more packages are available, so there are more alternative ways to build up the packages into ULD's. By scheduling breakdown operations just before buildup operations (as late as possible), packages do not need intermediate storage. If packages are directly transported from the breakdown process to the buildup process, storage-handling-in and storage-handling-out are avoided by direct internal transport and the required storage space is significantly reduced. On the other hand it should be noted that under the policy of scheduling as late as possible all float available for possible unforeseen circumstances is renounced at the outset.

Irrespective of the wisdom of these arguments, the model should be able to schedule as late as possible to reflect current practice. If the principles in the model cannot reflect current practice, the implementation of the model will fail. It will not only fail because computer planning differs from current manual planning, but also because users will not have any confidence in the model. In order to gain the confidence of the users, the model should be able to schedule as late as possible.

Like many other software systems SAS has the deficiency that it is not able to schedule activities as late as possible while taking resource restrictions into account. The deficiency is tackled by devising a transformation. All points of time that are part of the network and part of the resource availability specification are inverted. That means that the points of time (original earliest start times and original latest finish times and resource specification times) are mirrored in time by means of a reference time.

In formulae:
\[ x = \text{Set of all activities in the network.} \]
\[ E(x) = \text{Original earliest start for activity } x, x \in X. \]
\[ L(x) = \text{Original latest start for activity } x, x \in X. \]
\[ R = \text{Reference time.} \]
\[ E'(x) = \text{Mirrored earliest start for activity } x, x \in X. \]
\[ L'(x) = \text{Mirrored latest start for activity } x, x \in X. \]
\[ E'(x) := R - L(x). \]
\[ L'(x) := R - E(x). \]
\[ (E(x) \leq L(x), \text{ so } E'(x) \leq L'(x)). \]
\( Y \) = Set of resource specifications indices.

\( F_y(t) \) = Original resource specification \( y \) = Number of available resources in specification \( y \) at and from time \( t \) until next specification time, \( y \in Y \).

\( T_y(t) \) = Mirrored resource specification \( y \) = Number of available resources in specification \( y \) from previous specification time until and at time \( t \), \( y \in Y \).

\[ T_y(R - t) := F_y(t). \] (3)

After inversion the activities are as usual scheduled as early as possible. After scheduling all inverted points of time are mirrored once again by means of the same reference time. In this way a schedule is derived in which activities are scheduled as late as possible while resource restrictions are taken into account. The transformation is implemented in SAS.

The consequences for resources of latest scheduling will be presented with the help of graphs. The horizontal scale represents time and the vertical scale the amount of resources scheduled. A bar indicates the number of scheduled resources at a moment. A line represents the available number of resources.

Figure 5 depicts the resource profile for an arbitrary period. Activities are scheduled as late as possible without taking resource restrictions into account. As can be seen in Figure 5, resource shortages appear many times, but surpluses occur at other times.

Figure 6 shows the resource profile if activities for the same period as in Figure 5 are scheduled as late as possible while resource restrictions are taken into account. In Figure 6 no resource shortages occur at all.

The effect of the transformation (temporary time inversion) when activating the resource restrictions becomes evident from a comparison of Figures 5 and 6. In both figures activities are scheduled as late as possible. In Figure 6 activities are sometimes forced to finish earlier than their latest finish because of the active resource restrictions. The usage of resources is levelled under the availability line from the right to the left.

---

**Figure 5.** Resource profile for latest scheduling without active resource restrictions. Vertically the scheduled numbers of manpower teams are indicated by bars, the available numbers of resources by lines; horizontally time is represented in quarters of an hour.

**Figure 6.** Resource profile for latest scheduling with active resource restrictions. Vertically the scheduled numbers of manpower teams are indicated by bars, resources by lines; horizontally time is represented in quarters of an hour; leveling takes place from right to left.
left, i.e., shortages that occur at a given time in Figure 5 are scheduled at an earlier time in Figure 6.

The program includes a tracing option: activities that are scheduled when resource shortages occur can be listed. These activities can be considered as 'potential originators' of resource shortages. A unique identification indicates the particular ULD and the flight to which it belongs. By means of this option ULD flows can be adjusted systematically.

For example specific KLM flight origin stations can be asked to try and diminish their breakdown fraction. They are urged to build up their outgoing ULD's economically within the existing limited freedom for 'selective loading' and combine complete ULD's for connecting flights that can skip the breakdown and buildup processes at Schiphol Airport. The decrease of the breakdown fraction thus prevents foreseeable capacity problems in the hub.

7. Conclusions

The application of quantitative methods in airfreight operations is most suitable for medium term goodsflow control (medium term defined as: between two days and six months). In the long term goodsflow control, a lot of factors are dependent on the strategic policy of the company. It is hard to define and to value the variables associated with those 'soft' factors. For the short term goodsflow control, the number of factors is extremely large and a model would become unwieldy.

Improved control of the goodsflow can increase productivity. By adapting resource availability or airfreight characteristics to the resource profiles generated by the model, fewer operations are carried out late (i.e., effectiveness increases). At the same time the utilization of resources is augmented (i.e., efficiency increases).

The top-down simulation approach at the aggregation level of ULD's, together with the application of statistical techniques, reduces the complexity of the model without damaging the accuracy, reliability and validity of the results. Moreover it has the advantage that less data have to be processed.

Because the scale of the problem is too big, optimization in scheduling is impossible. Decision rules that generate non-optimal but satisfying results are needed, e.g., the rule to schedule operations as late as possible and the rule to level resource usage backward in the time dimension to accommodate arbitrarily preset capacity levels. Extra resources may be needed, even if violation of resource restrictions could have been avoided if other decision rules were used.

The big advantage of translating the scheduling problem into a project instead of a job shop is that the scheduling of operations can be carried out by standard software for network planning under resource restrictions. With KLM, the standard statistical and OR software is SAS.

Flexibility of the scheduling by the SAS software for network planning can be enlarged by applying some transformations: split of activities and inversion of time restrictions.

By application of standard software no queues for waiting operations have to be modelled and no decision rules have to be developed. Rules for levelling resources are incorporated in the standard software, cf. [3].

This paper has described the first step in an ongoing case study. The actual model should be improved in many ways: it should be refined and its scope should be increased. It will take some years before an empirical evaluation of the results of computer planning of the breakdown and buildup processes at the hub, as compared to manual planning, can be reported.

The study however convinced management in KLM of the usefulness of applying quantitative techniques in goodsflow control. Management is therefore determined to go ahead. And as is generally accepted, commitment of management is one of the key factors for successful application of quantitative methods.

References


