SIMULATING DYNAMIC DEPENDENCIES AND BLOCKAGES IN IN-PLANT MILK-RUN TRAFFIC SYSTEMS

Tobias Staab, Eva Klenk and Willibald A. Günthner
Lehrstuhl für Fördertechnik Materialfluss Logistik
Technische Universität München
85748 Garching bei München, Germany
E-mail: staab@fml.mw.tum.de, klenk@fml.mw.tum.de, guenthner@fml.mw.tum.de

KEYWORDS
In-plant milk-run; discrete-event simulation; blockages; in-plant road systems.

ABSTRACT
To provide efficient and flexible material transport, in-plant milk-run systems are used by many companies. However, the large number of influencing factors makes in-plant milk-run systems complex, while dependencies and blockages, e.g. overtaking or stopping, require a dynamic approach to examination. Thus we developed a generic simulation model focusing on traffic situations in in-plant road systems, and defined meaningful performance figures, e.g. capacity efficiency and cycle time. The application to a system of industrial practice enabled us to derive recommendations for planning in-plant milk-run systems. Space for overtaking, the number of trains passing or supplying a stop as well as the directions a stop is approached from were found to be main influencing factors.

MOTIVATION
In-plant milk-runs are a transport concept used for in-plant material supply from a storage zone to production, especially in the automotive industry. Since they are supposed to be efficient, safe, “lean” and flexible, especially for small batch sizes, many companies are working on implementing, re-engineering or standardizing their milk-run systems at present (Dreher et al. 2009).

In-plant milk-run systems supply different points of use with a variety of goods in one run (Takeda 1996). Typically, the milk-run “trains” use fixed routes. In some cases, they are operating on a fixed schedule. The system therefore is comparable to a bus system in public transportation. Depending on the size of the system, several milk-run trains may be operating from the same material source. Most often, milk-runs are tugger trains consisting of a tugger and three to five trailers. A typical milk-run tour consists of loading requested units of material onto trailers at a material source, travelling to the production area, stopping at different points, providing material, picking up empty load units and unloading all empty load units at the end of each tour (cf. Figure 1).

Figure 1: Typical Milk-Run Process

Milk-run systems are complex and dynamic. Several types of materials and types of load units are provided with a number of trains on a number of routes to various points of use. Usually, the number of load units required varies over time and is uncertain or known only a short time in advance. Moreover, in most milk-run systems driveways and resources, e.g. loading stations are used by more than one train. This may result in dependencies and blockages between individual trains, e.g. caused by overtaking or stopping vehicles.

As milk-runs are most often used in just-in-time systems with provision of small lot sizes in high frequency and low buffers of inventory at the points of use, a reliable and stable process is critical for success (Klug 2012). Defects in the process, e.g. late delivery or delivery to the wrong point of use may directly result in a production standstill and therefore high excess cost. Planners are challenged with the task of designing a milk-run system where a defined service level of provision is ensured and high efficiency (high capacity and time utilization) is achieved at the same time.

“Statically” dimensioning a milk-run system, i.e. deciding on which points of use are to be combined into one route and in which order is already a “hard” problem, as the underlying problem is similar to a vehicle routing problem with some added characteristics (Toth and Vigo 2002). The dynamic effects, uncertainties and blockages described above still cannot
be modelled adequately with such an approach, but may cause significant problems when operating a milk-run system.

Simulation models can be used to examine dynamic systems on a high level of detail (Schenk et al. 2008). Therefore, the paper on hand presents an event-discrete simulation model to model and analyze generic milk-run systems in a planning phase. Comparable models at present sporadically exist in some companies, but are specific to their individual projects and processes and unavailable to the public.

Our model focuses in detail on modelling the physical handling steps performed in typical milk-run systems (e.g. loading, travel, provision) and simulation of the resulting traffic and supply situation for any set of routes and various types of material requisition. The first requirement is a simple and adaptive structure. This enables planners to examine and compare various milk-run systems and to model and analyse different sets of routes quickly and simply. Secondly, all important dependencies between trains are to be modelled. Thus, possible bottlenecks and problems can be identified and solved in advance and better solutions found, e.g. by varying the routes. As a third requirement, requisition notes have to be simulated adequately as the reaction of the system to deviations of material requisition is to be considered. Furthermore, we aim to derive some universal recommendations for designing milk-run systems from our experiments with the model.

METHOD
Based on an analysis of physical handling steps performed in typical milk-run systems and of the requirements, we will identify relevant elements the reviewed systems consist of and focus on the non-specific elements. Afterwards, we will present the elements, their behaviour and interfaces we designed to establish a modular structure. As a result of the requirements as well as of the system behaviour we will define meaningful performance figures. We will apply the verified and validated simulation model to an example from industrial practice. Finally, we will derive recommendations to improve planning based on experiments using four scenarios.

ELEMENTS AND MODEL STRUCTURE
The increasing use of in-plant milk-run systems for production supply and the lack of established standardized concepts (Günthner et al. 2012) lead as a direct result to the application of various individual systems with their specific structure. Nevertheless, our simulation model shall likewise be applicable as generally as possible and with as few modifications as possible. For this reason, we chose a modular structure for our model. An analysis of existent milk-run systems shows that many elements appear in all milk-run systems and with little variation (Klenk et al. 2012):

- points of use
- stops
- driveways
- turns
- crossroads
- tuggers
- trailers
- load units

Hence, these objects are defined as object classes. These classes are modular and can be adapted with low effort to the respective system by a set of properties. They are sufficient to simulate a generic in-plant road system.

Some elements however, e.g. the loading areas of storages and the removal area for empties are system-specific, since they are influenced by a number of various factors, e.g. the type of trailers used, the size of load units and the space available. These elements must be modelled individually, but can be linked to the system using pre-defined interfaces. To permit an application to various systems, the interfaces between the elements of the road system mentioned above and the individual elements are reduced to a minimum.

Besides, the flow of information is also system-specific, depending on factors such as the available IT-infrastructure and the control system. Thus, the material booking is the only interface between the supply of load units and the stock management which is creating transfer orders.

As in-plant milk-run systems significantly differ from other transport concepts concerning the usage of the road system because of dependencies between trains, we will focus on modelling the traffic and road system. Thus, general applicability of the simulation model is ensured (cf. Figure 2).

![Figure 2: Elements and Interfaces in an In-Plant Milk-Run System](image-url)

MODELLING INTERDEPENDENCIES AND BLOCKAGES
Compared to production supply with autonomous participants, i.e. forklifts, the milk-run systems reviewed are liable to a fixed routing often combined with a fixed schedule. Each vehicle follows its determined route and never deviates from it, not even as a reaction to events like blockages on driveways. Based on this condition, the system cannot react to critical states, which makes
the dependencies between vehicles an object of central interest.

First of all, a control mechanism has to be installed at crossroads to regulate the right of way as well as the correct turn-off. As a restriction caused by the limited space in crossroads in reality, the maximum amount of simultaneously turning vehicles is set to two. Additionally, in case of more than two vehicles or incompatible turnings (crossing routes), waiting vehicles are given the right of way afterwards following the first-come-first-served rule.

The process of supplying transported material is the most time-consuming physical handling step and therefore liable to cause holdups or blockages. It is thus modelled in detail as any stopping vehicle forces potential successors to either wait or overtake, depending on the available space and whether the successor itself carries along material for the same stop. Steps included are the paths of the driver between tugger, trailer and hand-over places as well as the handling of load units. These steps are repeated for every load unit that is to be supplied. In addition, the simulation model also takes into account the lane the hand-over place is located on and the stopping position of the vehicle. As a result there are three different ways for the driver to reach the hand-over place (cf. Figure 3). The distances that are to be covered per load unit vary between five and eleven times the length of a trailer, not accounting for the breadth of the driveway. This variety of the distance and thus the resulting supplying time make it clear why this process is not simply estimated but calculated in detail: load peaks of the whole system shall not be levelled out by using mean values but shall be measured precisely. To reach this level of detail, all steps are modelled using Methods Time Measurement (MTM).

The distances that are to be covered per load unit vary between five and eleven times the length of a trailer, not accounting for the breadth of the driveway. This variety of the distance and thus the resulting supplying time make it clear why this process is not simply estimated but calculated in detail: load peaks of the whole system shall not be levelled out by using mean values but shall be measured precisely. To reach this level of detail, all steps are modelled using Methods Time Measurement (MTM).

For a better approximation of the real system, the influence of other disturbing factors can also be analysed. At first, some of the driveways are possibly used by other vehicles as well. As a result, the strain of the road system is likely to be increased, which is shown by the number of holdups and queue times, e.g. at crossroads. Additionally, driveways can be blocked by randomly appearing temporary obstacles, e.g. load units on lanes. Similar to stopping vehicles, obstacles also force oncoming traffic to circumnavigate if possible, identically to the overtaking process (cf. Figure 4), including the influencing factors mentioned.

Besides the crossroads control, supplying process, overtaking and disturbances, the loading in the storage zone is another source of dependencies because in general the number of loading points is lower than the number of vehicles. Thus, although the routes are phased in general, delayed tours directly affect successors because of an occupied loading point at the wrong time. On the contrary, vehicles arriving too early have to wait for their next departure in case of a schedule, preferably in a separate parking area to avoid the blockage of driveways. As these dependencies are subject to the system-specific layout and control mode they cannot be described in general terms.

As well as all kinds of dependencies the dynamic reaction of the system to a deviation of material requisition cannot easily be measured by analytic methods because of the large number of influencing factors such as the production program, the amount of pieces per load unit and the stocks in circulation. Hence, two mechanisms to create the material requisitions are provided corresponding to two levels of accuracy of the data available.
If the material requisition is known exactly in terms of time and amount of load units, original data can be used, e.g. a protocol of real Kanban orders. However, this data is seldom available if the system is not already in an advanced state of planning. Hence the second mechanism to simulate material requisitions is based on the average amount of load units per hour and creates periodic single material requisitions. As a result artificial material requisition peaks are generated by the random initial time of material requisitions of every material and the average time between two material requisitions. Both mechanisms can be combined for every combination of material and point of use. This allows widespread application of the simulation model to systems with parts in different states of planning.

PERFORMANCE FIGURES
To examine routing, supply, a possible schedule and the strain of the road system, a set of performance figures is necessary. Based on this standardized data, further examination is allowed to gain insight into the dependencies of the system and to derive recommendations to improve planning. Thus, the following figures are attained per route:

- As the timely supply of material is a central aim of logistics, the service level of a route is measured based on a default replenishment time and defined as the ratio of load units supplied in time and the number of load units transported in total.
- The capacity efficiency per tour measures the share of capacity occupied by load units. As the simulation model does not distinguish between differently sized load units, a larger capacity efficiency directly leads to more supply activities. On the contrary, if the size was considered, the replacement of a small load unit with a larger one would increase the capacity efficiency without changing the number of supply activities.
- Additionally, absolute delays in supply are measured for further clustering and analysis. Both the number and duration are important as a large number of low delays possibly does not affect the stability of production supply in the same way a few high delays could do.
- As soon as the stock at a stop reaches zero, the simulation aborts because of insufficient supply and logs the material and affected stop.
- To monitor the schedule, the effective time of departure is measured and compared to the reference value.
- The cycle time is defined as the time between the start of every tour and the arrival back at the loading area. It contains the constant time (per route) needed for travel as well as a variable part mainly consisting of supplying and waiting.
- As a direct result the temporal efficiency per cycle can be defined as the cycle time in reference to the scheduled cycle time.
- For every milk-run the total times of all activities such as driving, supplying, waiting, removal of empty load units and overtaking are measured separately for further analysis.
- Depending on the characteristics of the system-specific elements (cf. Figure 2) more performance figures can be necessary to rate the system, e.g. a protocol of the activities of a substitute train or the efficiency of a stowing forklift.

The strain of the road system can be rated by the following performance figures attained for every section of driveways:

- As waiting affects the traffic flow by holdups and resulting delays, the total number of waiting vehicles over time is counted on every driveway.
- In addition, overtaking and circumnavigation of obstacles as sources of waiting concerning oncoming traffic are counted over time per section.

APPLICATION TO INDUSTRIAL PRACTICE
After verification and validation of the simulation model, it was applied to an in-plant milk-run system designed for a plant of Brose CZ spol. s r.o. Simultaneous with planning, the current state was modelled. The system includes 3 production lines with a material need of over 300 load units per hour and a road system with more than 1,200 m of driveways and 20 crossroads over an area of 35.000 m².

To enable experiments, system-specific elements had to be added:

- As there are two types of milk-run systems to be analysed simultaneously, resulting from separated storages for small loads carriers (SLC) and large carriers (LC), there are two different loading processes:
  The loading of SLC is based on an automated “drive-through” concept with flow rack trailers (Dewitz et al. 2012). Each train is assigned to a certain route. When a train arrives at the SLC storage, it is assigned to one of four available loading stations, where loading takes place automatically.
  The LC are stowed by a forklift onto E-frame trailers. To reduce the time needed for loading, only the tugger is assigned to a route, while empty trailers are stowed for the next tour on the next route in the meantime and are changed between routes in the storage zone.
- As the loading stations of the SLC trains are filled automatically, trains arriving too early have to wait until their load is prepared. This requires a parking area adjustable to the number of trains to avoid a blockade of driveways and to generate the correct order of trains given by the loading points.
- To compensate for delays, substitute SLC trains can be used, which also requires a parking area for waiting in the meantime.
• Both types of trains use the same empties removal area. As the set of trailers is simply changed to one cleared before, the removal of empties is decoupled from the stopping time of the train.
• The information flow follows the second of the mechanisms given above using mean intervals between two material requisitions.

A simple heuristic was used for route-building: the material requisitions of machines and workplaces were consolidated to blocks bordered by driveways or space not belonging to the reviewed production lines. Each block constitutes the material requirement of an allocated stop. These stops are assigned to routes such that adjacent stops are on the same route and the total material requisition is lower than a defined maximum number of units. Simultaneously, the strain of the road system is considered: we tried for an even strain to avoid bottlenecks by equating the number of passing trains as far as possible for every section. Concerning the given system, 6 routes for SLC and 2 routes for LC are necessary with a constant scheduled cycle time per type.

EXPERIMENTS AND RESULTS
With the elements described above, the road system as well as the system-specific areas and processes could be modelled to an adequate level of accuracy. For the following experiments four scenarios were created to allow the analysis of the system and of influencing factors.

First, the system was parameterized not considering any disturbances to test routing and schedule and to use the resulting level of influences between the vehicles as a base for further experiments. Additionally a substitute SLC train was available to absorb delays. The total times for waiting, driving and supplying revealed the expected high influence of the supplying activities. i.e., the average share of supplying time was at about 75% for SLC. The total time for waiting was volatile over the routes. To examine this, cycle time and efficiency of a route rather unimpaired by waiting were examined (cf. Figure 5). While the capacity efficiency directly follows the material requisition, the cycle time mainly consists of the constant driving time and supplying time depending on the capacity efficiency. Thus the similarity of both graphs reveals the influence of the supplying activities.

On the other hand, a high share of waiting disturbs this similarity by adding waiting as a new relevant time component resulting from dynamic dependencies. To clarify this correlation, the ratio of the capacity efficiency and the temporal efficiency per cycle can be used as a similarity measure (cf. Equation (1)). This ratio can be interpreted as the resulting performance of the system \( P_{\text{res}} \) relative to the planned performance \( P_{\text{plan}} \):

\[
\eta_{\text{cap}} = \frac{\text{nsupp}}{\text{N}_0} \quad \eta_{\text{temp}} = \frac{\text{T}_{\text{sched}}}{\text{N}_0 \cdot \text{t}_{\text{cycle}}} \quad \eta_{\text{res}} = \frac{P_{\text{res}}}{P_{\text{plan}}} \quad \eta_{\text{temp}} = \frac{\text{T}_{\text{cycle}}}{\text{N}_0} \quad \eta_{\text{temp}} = \frac{P_{\text{res}}}{P_{\text{plan}}} \quad (1)
\]

Figure 5: Similarity of Capacity Efficiency and Cycle Time
To estimate the impact of waiting, this ratio is determined for a route unimpaired by waiting (cf. Figure 6) as well as for a second one with a significantly higher share of waiting. A higher waiting time increases the variability of each graph, whereas the conformity of capacity efficiency and cycle time would cause a constant ratio (cf. Figure 6). However, the absolute value of between 0.7 and 0.9 is a matter of the planned performance \( P_{\text{plan}} \) influenced by different system-specific parameters such as the total capacity of the trains and the length of each single route. If differently sized load units are considered, Equation (1) must be adapted, as in this case the capacity efficiency is not directly scaling with the number of supply activities.

Over all routes, the cycle time maxima as well as the share of tours with an exceeded planned cycle time reveal that the observance of the schedule varies. As a result, the substitute SLC train is activated to compensate delays on different routes. In this case, the definition of a substitute temporal efficiency as the ratio of active time and total simulation time seems to be useful. Caused by the varying observance of the schedule, the substitute temporal efficiency is between 13% and 20%. When comparing the route the substitute train had to take significantly often with the most strained sections, a clear correlation could be found.

Figure 6: “Similarity” of Capacity Efficiency and Temporal Efficiency
For the second experiment, the substitute SLC train was removed to compare the expected negative effect on the system to the saving of a train and driver held on standby. The results of the first experiment indicated that the lack of a substitute train cannot be compensated if the substitute temporal efficiency is
establish the supply by a single train, which is impossible in the reviewed system because of the different processes for SLC and LC. Furthermore the allocation of a stopping point to more routes would lower the average supply time per stop and tour and simultaneously the waiting time per follower. If the space available is sufficient for two trains and the drivers who handle load units on each trains’ right side, an approach from different directions could be utilized. However, the effect would only be appropriate if there are no further trains passing through because overtaking would be aggravated at the same time. The supplying process is much faster if the driver does not have to circle the train to reach the opposite side of the driveway (cf. Figure 3). Thus, matching both the approaching direction and the side of the hand-over area is another way to lower the strain of the road network.

CONCLUSION AND OUTLOOK

Due to the complexity and size of the system, computer-aided methods such as discrete-event simulation allow the analysis of the material supply of in-plant milk-run systems. As a result of the large variety of these systems the design of certain elements of the model is subject to the specific type of the system.

Thus, the general application was solved by a modular structure of the elements necessary for modelling an in-plant road network. Certain elements, such as areas for loading and removal of empties, cannot be pictured generally. Hence, lean interfaces were designed to allow the simple integration of elements with a system-specific architecture. The information flow contained in the model described above is a Kanban system and allows both the usage of real data as well as the application of mean material requisitions to model material requisition. Like the elements of the material flow, this material requisition mechanism can easily be replaced using the described interfaces. Generic elements of milk-run systems (pathways, crossroads, trains, …) were modelled and focused on picturing all relevant dependencies between the vehicles passing the system, such as behaviour at crossroads, supplying, overtaking and disturbances.

Subsequently a model of milk-run systems that are currently realized at Brose CZ spol. s r.o. was built, containing eight routes for two types of load units with different trailers and loading processes. In a series of experiments, routes as well as the influence of disturbances were tested.

As a result, the influence of the number of load units supplied and of waiting processes could be confirmed. Since they allow the simultaneous analysis of both, capacity efficiency as well as cycle time appear to be important performance figures. Based on this, the ratio of capacity efficiency and temporal efficiency can be used as a quantitative “similarity” measure for the
effect of waiting on a particular route. This ratio can also be interpreted as the ratio of real and planned performance. In addition, the strain of the road system is revealed by the number of overtaking and waiting processes per section. One main influencing factor is trains supplying at stops.

Important recommendations concerning the planning of in-plant milk-run systems were derived from the results. If a stop is approached from both directions, additional trains passing should be avoided. Two-lane driveways and space to overtake are important to lower the number of holdups and improve the traffic flow. In general, the strain of the road network is reduced by a low number of trains passing or supplying the same stop. The number of load units per tour and stop is another main influencing parameter because it determines the supply time the trains needs at a stop.

Some additions still need to be made to the system: Overtaking and turning off at crossroads are not yet combined because of the complexity of new situations related to this, e.g. caused by the incompatibility of the directions the involved vehicles are heading for. In addition, each vehicle can at the same time only be overtaken by one successor, which causes waiting times that do not occur in reality. As the system reviewed in the paper on hand is still subject to planning, the current results are to be compared to recent data based on an advanced status.

Also, we are currently working on developing further modules for other loading concepts, to broaden the application of the model.

Additionally, further experiments will be carried out to examine the effects of one-way traffic and of splitting material requisition onto more and smaller stops.

REFERENCES

Dewitz, M.; S. Galka; W.A. Günthner. 2012. “Drive-Thru Loading Concept for In-Plant Milk Runs”. In: Proceedings of the XX International Conference on Material Handling, Constructions and Logistics (Belgrade, Oct.3-5). University of Belgrade Faculty of Mechanical Engineering, 237-242


Schenk, M.; J. Tolujew; T. Reggelin. 2008. “A Mesoscopic Approach to Modeling and Simulation of Logistics Networks”. In: Logistics and Supply Chain Management:


AUTHOR BIOGRAPHIES

TOBIAS STAAB studied Automotive and Combustion Engine Technology at the Technische Universität München (TUM). He obtained his degree in 2012 and has worked as a research assistant at the Institute for Materials Handling, Material Flow, Logistics (fnl) of TUM since 2013. Focus of his work is the simulation of supply chains.

EVA KLENK studied Industrial Engineering at the Karlsruhe Institute of Technology (KIT). Since 2009 she has worked as a research assistant at the Institute for Materials Handling, Material Flow, Logistics (fnl) at the Technische Universität München (TUM). Her main field of research is lean process design for automotive supply chains. Since 2012 she is leading the research group Process Design at the Institute fnl.

WILLIBALD A. GUENTHNER is professor and head of the Institute for Materials Handling, Material flow, Logistics (fnl) at the Technische Universität München (TUM). He is founder member and treasurer of the Wissenschaftliche Gesellschaft für technische Logistik e.V, deputy chairman of the scientific advisory board of the Bundesvereinigung Logistik (BVL) and member of the board of directors of the Society for Production and Logistics of the Association of German Engineers (VDI-GPL).