

## **EVENT-BASED RECOGNITION AND SOURCE IDENTIFICATION OF TRANSIENT TAILBACKS IN MANUFACTURING PLANTS**

Clemens Schwenke  
Thomas Wagner  
André Gellrich  
Klaus Kabitzsch

Institute of Applied Computer Science  
Dresden University of Technology  
01062 Dresden, GERMANY

### **ABSTRACT**

Automated material handling systems in complex manufacturing plants oftentimes exhibit, possibly transient, tailback phenomena, which reduce a production line's throughput. In order to identify origins and causes of observed tailbacks, historic event log data of loads passing certain waypoints has to be inspected. This paper introduces an approach to automatically carry out transient tailback recognition *and* cause identification. The approach is based on analysis of holding times and capacities of transport segments. As a result, complete lists of tailbacks and affected segments are provided. Each tailback consists of a reconstructed queue of loads waiting for preceding loads. For each tailback an initial cause event is determined. Additionally, identified tailbacks can be ranked by length or by impact on the transport delays. The developed demonstrator frees the user from time consuming visual inspection of log files by providing clearly represented complete tailback information instead.

### **1 INTRODUCTION**

The large amount of production steps in modern fabrication on the one hand, and the flexible use of automated equipment for a great variety of different products on the other hand, cause manufacturers to build and operate very complex production lines. These consist of many processing stations that are connected by interwoven networks of transport elements such as conveyor belts and handling devices.

These so-called automated material handling systems oftentimes exhibit tailback phenomena, which reduce a production line's throughput. Tailbacks are queues of loads waiting for being processed in front of a workstation. The reason is that workstations sometimes are incapable of processing the arriving loads fast enough. The resulting tailbacks can be transient phenomena because workstations can be overloaded or shut down only temporarily. Consequently, after overload or shutdown periods end, the tailbacks dissolve. In order to react to current tailbacks, the input rates of loads get reduced or modern material handling systems even adjust to undesirable situations and dissolve tailbacks by the use of, e.g., dispatching rules.

This is beneficial for the throughput, but since the tailbacks dissolved, potential, possibly transient, problems cannot be identified anymore. Therefore, in the past event data of material passing certain waypoints of the system has been collected in so called log files. But the task of identifying transient tailback phenomena is still carried out by humans reading many lines of log files. This paper introduces an approach to automatically carry out transient tailback recognition *and* cause identification.

As an exemplary use case, a compact set of event data of a car manufacturing line has been used to illustrate and validate the approach. The approach can be applied for analyzing any rail bound transport

system, which are common in the semiconductor industry. The remainder of the paper is structured as follows. Section 2 reviews briefly the related work in the field of analysis of event data. The exemplary use case is explained in Section 3. Section 4 describes the necessary steps for processing the data and shows an overall workflow of the approach. The results are demonstrated in Section 5. A conclusion and outlook are given in Section 6.

## 2 RELATED WORK

As mentioned above, the introduced approach, incorporates three basic methods of investigating recorded event data: state model building, descriptive statistics, and systematic backtracking of loads as a way of analytical cause identification. In order to apply the latter two methods, first the recorded event data has to be prepared. Here, this preparation is carried out by state model building. The result is a state model, extracted from the event data. It is a representation of all observed event sequences in one graph. This graph then is the basis for statistics identifying states with overload as manifestations of tailbacks. Then systematic backtracking to the cause events and states is carried out. In the following, related work is considered.

### 2.1 State Model Building

Automatic state model building requires event logs as input data and provides highly aggregated information about state changes in a system. The prerequisite is that recorded events can be understood as notifications of state changes of entities. Briefly, the main essence of state model building is the extraction of a graph out of traces of events.

In the case of material flow systems, events are recorded when loads enter workstations where they are processed, or when they enter conveyor segments where they are transported to a succeeding workstation. In the extracted graph, nodes represent states of loads being in a certain station or transport segment, and edges represent the transfer of a load into another station or transport segment. For this kind of event data, an event always indicates that a load entered a work station or transport segment. The use of discrete state models describing a system's (or device's) behavior as a sequence of possible steps was studied successfully before (Kemper and Tepper 2005).

On the one hand, event logs of machines or transport systems are used for state model building, for example, in semiconductor industry or in logistics applications (Gellrich et al. 2010). In order to enrich a pure state model with more information, Vasyutynskyy suggested the combination of state model building with calculation of performance indicators, such as loads' and stations' overall throughput times, holding times, and inter arrival times on states. Consequently the result is called an *extended* state model (Vasyutynskyy et al. 2010).

On the other hand, event traces of simulations can be examined by applying the method of state model building as well. In contrast, common analysis of simulation results oftentimes is limited to acquisition of indicators like average throughput or maximal transport time. But averaged statistics allow only a rough estimate of the overall system performance, not considering its internal behavior and causes for observed specification violations. A detailed analysis of possible problems like bottlenecks or control logic faults is often possible only by visual inspection of simulation results based on log data or in the visual interface of the simulators. Consequently, Kemper et al. developed several approaches for event log analysis. A visual inspection approach was developed for clustering the time intervals to differentiate between normal and abnormal dynamic behavior. Changes in logs that point to significant changes of the system's behavior had to be detected (Kemper 2007).

Besides, state models are useful to monitor or identify business processes (Agrawal et al. 1998). Van der Aalst et al. (2007) used state model building as a method for analyzing business processes, where events are generated when certain work steps begin and end. In so-called process mining, models of business processes shall be recovered or checked. Additionally, the relevance of business steps can be evaluated and performance indicators are calculated based on event logs. The main problems are to recover ad-

equate models and to identify relevant process steps, since the log data of business processes, involving humans and external events, oftentimes contain non-deterministic fractions. The resulting models then have to be mostly analyzed manually, sometimes including a few automatically calculated performance parameters if applicable. Compared to extracted models of business processes, the extracted models of logistic and manufacturing applications are more deterministic but contain many more states, so that sophisticated tailored analysis approaches are necessary to detect and explain unwanted phenomena, such as transient delays (Gellrich et al. 2011) or changing reject occurrence (Shanthikumar et al. 2007).

An example of process mining is the analysis of historic event data of maintenance activities with the purpose of identification and prediction of equipment faults and necessary maintenance service. Modeling and analysis of recorded event series was based on state model building. Indicators, e.g. fault rate and duration of repairs, determine the operational availability and expenses of a device. The time intervals between events are important (dynamic) factors affecting these indicators. Therefore, the investigation of time intervals influencing the indicators is the main focus in that case study (Schwenke et al. 2011).

## **2.2     Statistic Approaches**

The flow of indistinguishable items, e.g., bottles, is generally analyzed by applying averaged statistics (Struss and Ertl 2009). In contrast, in automated factories, and especially in semiconductor fabrication, work pieces nowadays have identification numbers. This allows us to distinguish, trace, and evaluate the items' transport performance individually (Gellrich et al. 2010).

A survey by Nazzal and El-Nashar (2007) concluded that hardly any literature has been found on the subject of failures of conveyor segments that might cause congestion and deadlocks in the system, whether in semiconductor simulation studies or in analytical models. Therefore, they stated that more research work will be needed, e.g., in the areas of modeling closed-loop conveyor material handling systems based on analytic methods, development of analytical and simulation models which consider the joint optimization of several design problems such as those of conveyor systems and of other material handling systems like stockers or under-track storage units.

As expected, more-sophisticated work was done subsequently in the area of analysis of semiconductor transport systems. The influence of available overhead hoist vehicles in a transport system on delivery times as well as load and unload times of wafers was examined by Govind et al. (2011).

A summarized statistical comparison of scheduling strategies and load cases was carried out by Zhou and Rose (2011). Here the work-in-progress fluctuation was in focus and the influence of different scheduling and load cases was evaluated by calculating summarized statistics on cycle times and on the overall share of work-in-progress pieces (waiting for operation) in ratio to all wafers.

Completely based on statistics extracted from historical lot data traces, Mosinski et al. (2011) forecasted delivery times of wafers. In this work, the time estimation is based on a statistical evaluation of the duration between start and target operations from historical data. The time interval differs for different manufactured products. Therefore, wafers were grouped into classes based on product characteristics, priority, and tardiness. For calculating the prediction, the method uses the current wafer position on the route and its corresponding class, so that arrival times at the target are determined by adding the previously defined duration to the current time.

Work on selecting the appropriate statistical techniques depending on the available data was contributed by Kleijnen (1999). In general, the aim of applying the suggested hypothesis tests is to validate simulation models but not to backtrack to the origin of tailbacks.

All this work is essential for determining and estimating performance indicators like throughput and delivery times, but these papers do not investigate interrelations and phenomena of wafer interaction, such as transient tailbacks, in the greatest possible detail yet.

### 2.3 Analytical Cause Identification

As mentioned, there is little work on identifying and evaluating tailbacks of work pieces in rail bound transport systems for distinguishable piece goods. But, there is work addressing related issues, e.g., tracing individual work pieces with the aim of explaining outliers (Kemper and Tepper 2007).

Analysis of internal logistic systems based on event logs with the aim of identification of anomalies in the behavior of these systems was carried out by Vasyutynskyy et al. (2011). But in this work quantification of anomalies was sufficient only for the use case of designing airport facilities.

A very holistic and rich in detail approach involving modeling and simulating a semiconductor factory was provided by Scholl et al. (2010) with the aim of forecasting delivery times of wafers. This simulation model was mainly developed in order to forecast overall delivery times but it would allow insight into the interaction of wafers because it provides actual process routes, equipment behavior, lot size, available processing modules, process dedication at the equipment level, equipment downs at the main-frame level, lot release strategy, and dispatch rules.

There is work about combining simulation with automatic data log analysis with the aim of validating the simulation, e.g., by Verbraeck and Versteegt (2001) and Saanen (2003). Focusing on the test of control systems Versteegt and Verbraeck (2002) used track data inspection methods in order to check the correct behavior of modeled transport systems. Advanced statistical techniques that integrate cause and effect analysis have been studied by Sargent (2004). Inspection of historic data has been applied in order to build a model and determine whether the model behaves as the real system does.

Most approaches focus on using simulations for the design of future facilities by, e.g., plants or airports. But they do not exhibit tailback reconstruction with assignment of individual sources to each detected tailback involving data of existing manufacturing lines. In general it can be stated that most data log analysis approaches that are published either aim at calculating averaged statistics, e.g., average processing times or cycle times or that advanced log inspection methods are not described in greatest possible detail.

Nazzal and El-Nashar (2007) also stated, that efficient analytical models for traffic management in conveyor systems with capacity constraints should be developed. Therefore, the authors contribute an approach for this issue. Consequently, the tailback recognition approach introduced in this work goes beyond applying only statistical methods by combining them with detailed tracking of work pieces in order to determine which piece causes which to wait. This shall be understood as *one* step on the way towards the goals formulated by Nazzal and El-Nashar (2007).

## 3 USE CASE

The exemplary use case is a manufacturing plant consisting of work stations and conveyor belts representing transport elements. Events are recorded and stored into a log when work pieces, also referred to as loads, enter work stations where they are processed or when they enter conveyor segments where they are transported to a succeeding work station. For simplicity, workstations and conveyor segments simply will be called stations.

The event data has been made anonymous in this paper. The data contains a subset of trace data of loads in an automated production line and has been used for proving the approach. The analyzed data set consists of 10,261 events that indicate the transportation of 1,417 loads within a system of 27 conveying and working stations in a time frame of 16 hours.

Figure 1 shows the graph of the state model that was extracted from the event data by state model building, as described in Sections 2.1 and 4.2. In the left, at four source stations (ST41, ST42, ST43, ST44) loads enter the plant. At the last station at the right, (SINK), loads leave the plant. In this plant any load that enters one station first will be processed and released first (FIFO). Tailbacks manifest in the form of longer holding times of loads at stations due to load case changes or downtimes at stations. Anticipating results, in Figure 1 the states have been colorized according to their frequency in causing tailbacks. The detection and representation of tailbacks is described Section 4.

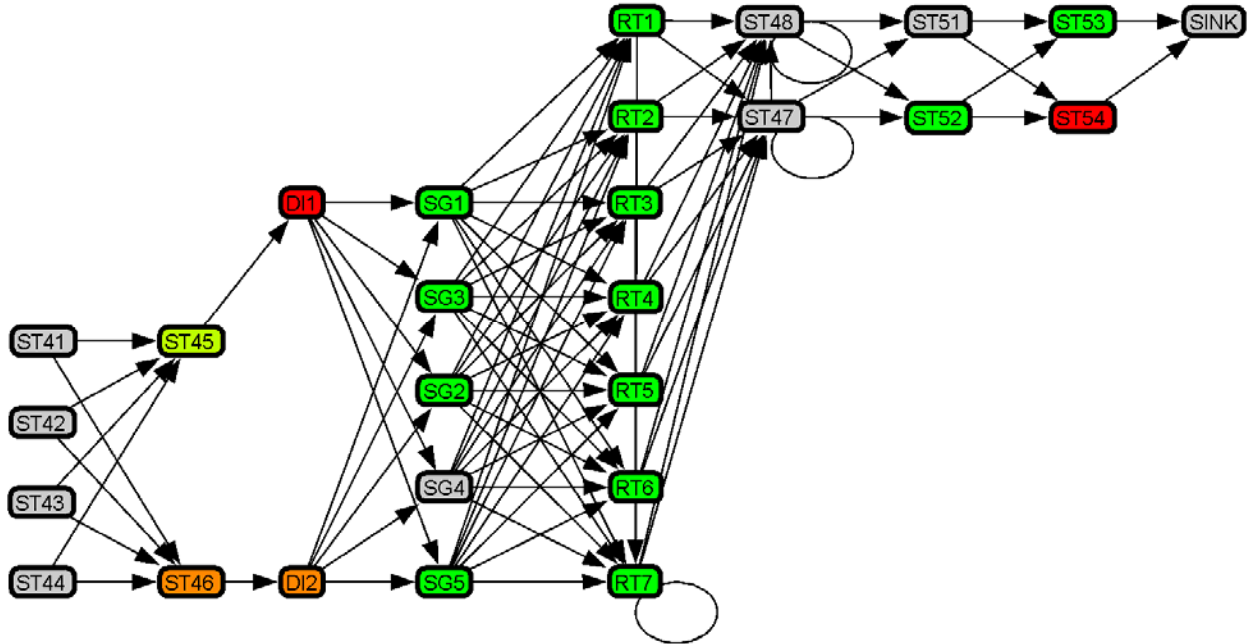


Figure 1: State model of manufacturing plant

## 4 DESCRIPTION OF THE APPROACH

In order to analyze tailbacks, two main tasks have to be carried out. First, points in time, when stations show overload, have to be found. Manifestations of overloads are longer holding times of loads in a station during a certain time frame. Second, the causes for these overloads have to be identified. These causes are difficult to find, because they are mostly located in stations that do not show manifestations of overload phenomena themselves. Analysis of tailbacks is based on the extraction of state models from event logs. Accordingly, the overall approach involves the following steps; cf. Figure 2.

### 4.1 Logging of Relevant Event Data

The approach uses historical load transport event data containing information about the time and location of a load. The location is pictured by the station the load was transported to or processed by. Consequently each log file entry is a triple consisting of a timestamp, load name, and station name. In order to collect useful data, logging of such events should be performed at each source, sink, and junction.

### 4.2 Building a State Model of the Transportation Process

Based on historical event data, a state model consisting of nodes and edges can be constructed. Nodes depict transport and working stations. Edges are directed and depict the observed material flow; cf. Figure 1. The building of state models is the basis for calculating aggregated information about each state, e.g., stations' average holding time. Average holding times have to be calculated for time frames and can be presented as time series. Average holding times per time frame are colored continuously by their length, so that red segments represent time frames of stations with extremely long load holding times; cf. Figure 3.

### 4.3 Identification of Temporarily Overloaded States

First, the overall average holding time of each state is calculated as the usual time a load is being processed or transported by a station. Then, each station's coefficient of variation, as the quotient of the standard deviation and average holding time, is calculated to identify stations which *sometimes* show very

long individual holding times of loads. The algorithm can be parameterized with a threshold for the coefficient of variation so that stations with relatively small standard deviations will be considered as acceptable and will not be investigated further.

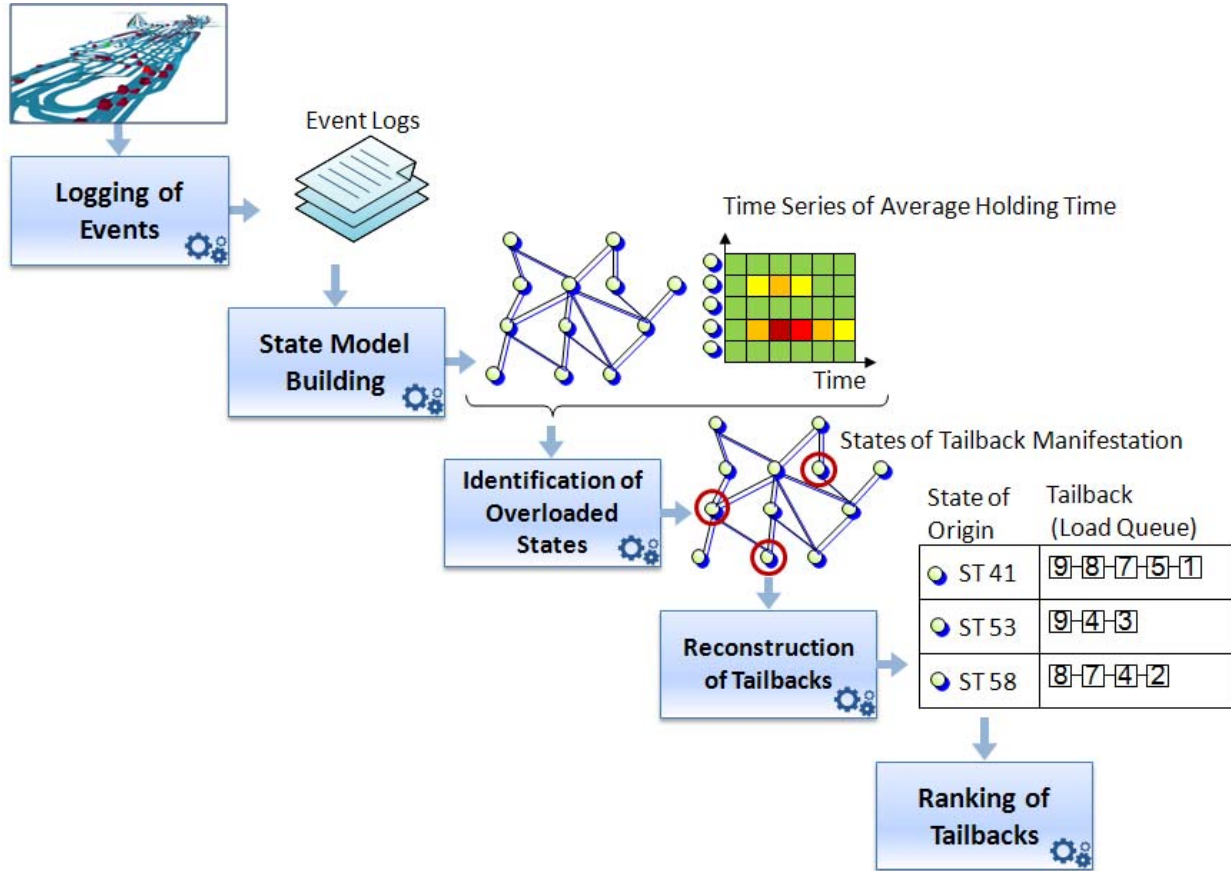


Figure 2: Workflow of tailback analysis

#### 4.4 Reconstruction of Tailbacks

Step 3 identifies stations where loads spent unusually long times being processed or, more likely, waiting for the successor stations to be capable of receiving these loads. Loads show a longer holding time in one station if the capacity of the succeeding station is reached or if the load was simply held up for other reasons that are not reflected by entries in the log file.

For evaluating whether a load's individual holding time is acceptable or rather too long, the algorithm can be parameterized with a threshold for acceptable holding times. A load's acceptable holding time should be set between a station's overall minimum holding time and the average holding time. Consequently, the parameter is called the acceptable holding time factor and is multiplied by a station's minimum holding time.

If it is detected that a station's capacity was reached, the reason for this can be assumed to be due to a succeeding station. On this succeeding station the same question has to be answered again. Is a load simply held up for other reasons or has the next succeeding station's capacity also been reached? The algorithm terminates if a load that is being held up was detected as a cause, although the succeeding station could have received it.

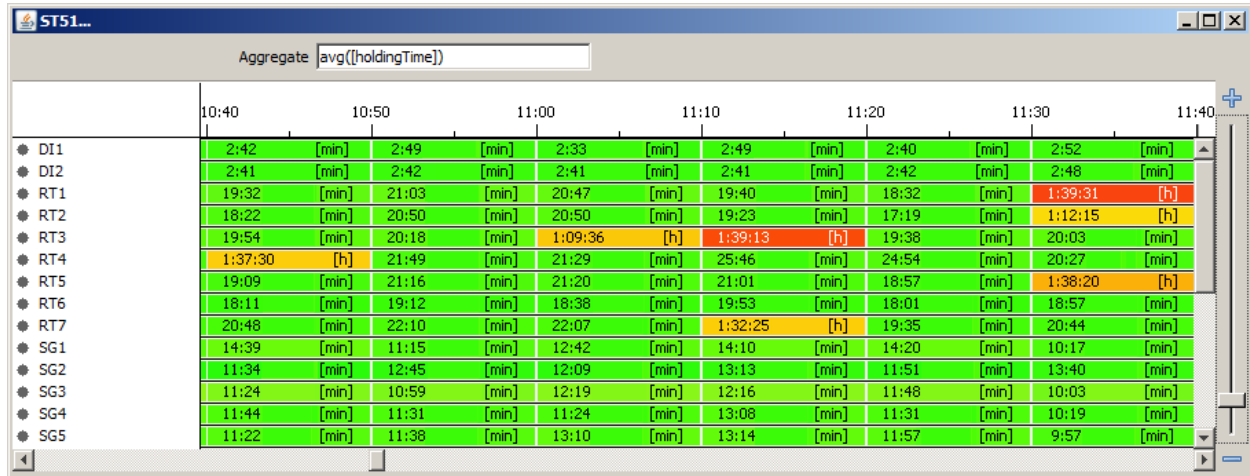


Figure 3: Average holding time of stations based on time frames

By carrying out this recursive test for each succeeding station, queues of loads waiting for preceding loads are reconstructed. The load the algorithm added last is the first load that was held up for unknown reasons. This load can be assigned to a definite entry of the log file. This entry is the sought-after cause event. This event points to the station that is the causing state or “source” of this tailback.

#### 4.5 Ranking of Tailbacks

Since the above-described reconstruction of tailbacks normally provides many tailbacks, they should be ranked. A ranking of tailbacks enables the user not only to find the most-severe tailbacks, but also to identify the stations that most-often cause tailbacks. The ranking can be based on the length of load queues or on tailbacks’ impact on the deviation of holding time on stations; cf. Figures 4, 5.

Table 1						
Table Diagram Scatter Plot Help						
Table view (900 rows)						
Index	Length	Cause	Events			longest HT
0	53	ST46	[7018@ST42@11:18:53 [h],	7024@ST41@11:20:05 [h],	7021@ST43@11:20:10 [h],	3:37:44 [h]
1	47	DI1	[7147@ST45@14:21:56 [h],	7223@ST42@14:32:31 [h],	7235@ST42@14:41:05 [h],	3:45:52 [h]
2	43	DI2	[7396@ST46@17:33:05 [h],	7381@ST46@17:34:29 [h],	7400@ST46@17:35:54 [h],	2:16:10 [h]
3	37	DI1	[7111@ST45@13:49:40 [h],	7172@ST44@13:50:14 [h],	7176@ST44@13:53:06 [h],	2:19:11 [h]
4	36	ST54	[7307@ST52@17:31:53 [h],	7311@ST52@17:33:15 [h],	7304@ST51@17:33:16 [h],	2:14:48 [h]
5	33	ST54	[6895@ST51@11:48:23 [h],	6842@ST51@11:49:45 [h],	6899@ST52@11:49:50 [h],	3:43:45 [h]
6	31	ST46	[7412@ST41@16:57:31 [h],	7416@ST41@17:00:23 [h],	7401@ST44@17:01:03 [h],	3:46:04 [h]

Figure 4: Tailback table, ranked by queue length of waiting loads

The information about tailbacks can also be used to color the graph of the state model of the manufacturing plant. By coloring the states according to their frequency in causing tailbacks the user gets a quick overview; cf. Figure 1.

Table 1						
Table view (900 rows)						
Index	Length	Cause	Events			longest HT
134	3	ST45	[7012@ST41@11:11:31 [h],	7016@ST41@11:14:22 [h],	7020@ST41@11:17:14 [h],	7:25:29 [h]
84	6	ST45	[7130@ST43@13:18:02 [h],	7135@ST42@13:18:48 [h],	7134@ST43@13:20:53 [h],	7:16:30 [h]
23	15	ST46	[6666@ST41@6:25:05 [h],	6664@ST42@6:25:15 [h],	6670@ST41@6:27:56 [h],	7:08:56 [h]
25	14	ST45	[7110@ST44@13:04:32 [h],	7121@ST41@13:05:17 [h],	7128@ST44@13:15:57 [h],	7:08:49 [h]
38	10	ST45	[7272@ST43@15:09:50 [h],	7284@ST43@15:18:24 [h],	7288@ST43@15:21:15 [h],	7:04:15 [h]
61	7	ST54	[6867@ST51@11:22:18 [h],	6695@ST51@11:26:24 [h],	6871@ST51@11:27:46 [h],	6:54:50 [h]
161	2	ST54	[6530@ST51@6:38:43 [h],	6537@ST51@6:45:32 [h],		6:41:09 [h]

Figure 5: Tailback table, ranked by impact on stations' average holding time

## 5 RESULTS

The approach provides a number of tailbacks with information about their length and number of involved loads. By spot samples of known tailbacks, the cause event identification in relation to the tailback manifestation was rechecked. The granularity of the algorithm's output can be modified by adjusting two parameters. The first parameter is the threshold for coefficient of variation of a station's average holding time. The second parameter is the threshold for acceptable holding times of loads on stations.

In order to investigate purposeful settings for the algorithm's parameters, different settings have been examined. First only one parameter was varied and the result is depicted in Figure 6. Then both parameters were varied and the results are shown as 3-dimensional bar charts in Figure 7.

### 5.1 Variation of Threshold for Coefficient of Variation

The variation of only one parameter, namely the threshold for acceptable coefficient of variation of the stations' average holding time, was varied in the range from 0.0 to 1.4 in steps of 0.1. The acceptable holding time for individual loads' holding times on succeeding states was fixed at 1.1 of the minimal holding time of a state; cf. Figure 6.

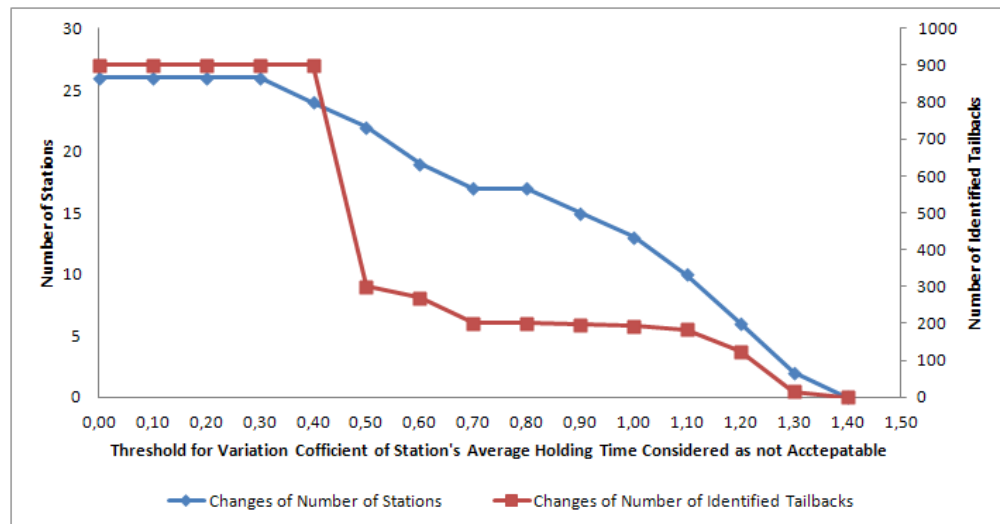


Figure 6: Number of identified stations and tailbacks in response to threshold for coefficient of variation

The result is that in the range from 0.0 to 0.4, the maximum number of tailbacks (900) is found, i.e., those tailbacks that are very short and resolve quickly are found. At 0.5 the number of affected stations

starts decreasing (22) but the number of identified tailbacks drops dramatically from 900 to 300. This is an indication for a good setting for the threshold. The number of identified tailbacks stays approximately the same until the threshold reaches 1.1. At 1.4, no tailbacks are identified anymore because the algorithm considers the (sometimes very big) variations of all stations as normal. It can be concluded that useful settings of the threshold for acceptable variation coefficient range between 0.5 and 1.1. In this range, the number of indentified tailbacks differs still. This is due to the fact that some tailbacks are short or their impact on delay on holding times is rather small. The user can adjust the algorithm to his needs.

## 5.2 Variation of Thresholds for Coefficient of Variation and Acceptable Holding Time

Both parameters of the algorithm were varied. The threshold for acceptable coefficient of variation was varied in a range from 0.3 to 1.2 in steps of 0.1 and the factor for the threshold of loads' acceptable holding times was varied from 1.9 to 1.1 in steps of 0.1. Consequently, the results for settings of 110 combinations of parameter values had to be calculated; cf. Figure 7.

As mentioned, the variation coefficient of stations' overall holding time separates acceptable stations from those that are considered suspicious and therefore deserve further investigation. The influence of the threshold for the acceptable variation coefficient is as described in Section 5.1. The higher the variation coefficient, the more tailbacks are ignored. The lower it is, the more tailbacks will be recognized; cf. Figure 7. For data sets with fewer tailback phenomena, lower thresholds should be used. For data with a high number of possibly stronger and weaker tailback effects in the same data set, higher thresholds should be used first in order to get an overview. Then step by step, smaller thresholds can be used to find less-important tailbacks.

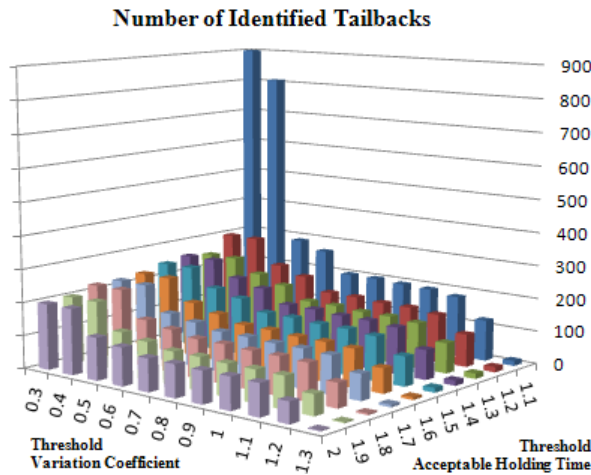


Figure 7: Number of identified tailbacks

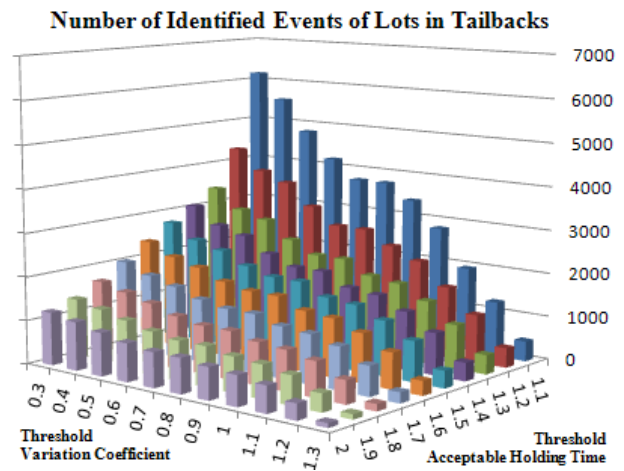


Figure 8: Number of corresponding events

The acceptable holding time factor influences when the reconstruction of load queues terminates because a load's individual holding time is considered acceptable. Therefore, the load was not waiting and no further cause event has to be searched for. A load's individual holding time is compared to the threshold, which is the product of a station's minimal holding time multiplied by this factor. If the load's individual holding time is bigger than this product, the threshold is reached. That is, the load's holding time is considered unacceptable and another preceding load on a succeeding station has to be examined. Otherwise, the algorithm terminates. In general, bigger threshold factors for the acceptable holding time set the algorithm to be more tolerant and result in shorter load queues; cf. Figure 8.

As depicted in Figure 7, the threshold of stations' coefficient of variation should be set rather low and the factor for the threshold of loads' acceptable holding times should be set rather high in order to find all

possible tailbacks; but such settings assign many events to tailbacks. This is not always purposeful. Consequently the user has to determine useful settings for his plant and use case.

## 6 CONCLUSIONS AND OUTLOOK

At present in industry, often-times time-consuming visual inspection of log files is the method of choice if transient tailback phenomena shall be investigated. One serious difficulty is the dispersion of the crucial events in the log file. Additionally, each tailback occurrence has to be looked at individually. Plus, by visual inspection the user is never sure if he has found *all* important tailbacks and *how* important they are.

The innovation of this approach is that all of these calamities are resolved by automatic tailback recognition and ranking. The user now can quickly find interesting or rather problematic stations as well as cause events. Besides this approach being feasible for analysis of a real plant's event data, it can also be applied for analysis of simulation, which is common for planning conveyor systems in logistics as well as in the semiconductor industry.

As a result of the work, a demonstrator was developed by the authors; cf. Figures 1 and Figures 3-5. Automatic tailback recognition, identification of cause events, and ranking frees the user from visual inspection, if he wants to find transient tailback phenomena. In addition, we provide two threshold values which can influence the granularity of the algorithm's output and therefore its calculation costs. This way, the algorithm can be adjusted to the user's purpose.

Further work will include the validation of this approach on bigger test data sets from real manufacturing plants, as well as simulations and the integration of the approach into an overall methodology of investigating factory event data.

## ACKNOWLEDGMENTS

This work was funded by the German Federal Ministry of Economics and Technology (BMWi), through the Alliance for Industrial Research (AIF) and the German Logistics Association (BVL) within the research project O-MFD (Online Material Flow Diagnosis) under the reference number 17343 BR. The authors are responsible for the content of this publication.

## REFERENCES

- Agrawal, R., D. Gunopulos, F. Leymann. 1998. "Mining process models from workflow logs". In *Proceedings of the Advances in Database Technology (EDBT'98), 6th International Conference on Extending Database Technology*, Valencia, Spain. *Lecture Notes in Computer Science*, vol. 1377, 469–483. Berlin: Springer-Verlag.
- Gellrich, A., V. Vasyutynskyy, K. Kabitzsch. 2010. "Using SECS/GEM Logs for Analyzing Internal Transport Processes in Semiconductor Equipment". In *Proceedings of the 10th European AEC/APC Conference*, Catania, Italy, 106–108.
- Gellrich, A., T. Wagner, V. Vasyutynskyy, K. Kabitzsch. 2011. "Modeling of Transport Times in Partly Observable Factory Logistic Systems based on Event Logs". In *Proceedings of the 16th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA 2011)*, Toulouse, France.
- Govind, N., T. M. Roeder, L. W. Schruben. 2011. "A Simulation-Based Closed Queueing Network Approximation of Semiconductor Automated Material Handling Systems". *IEEE Transactions on Semiconductor Manufacturing*, vol. 24, no. 1, 5–13.
- Kemper, P., C. Tepper. 2005. "Trace based analysis of process interaction models". In *Proceedings of the 2005 Winter Simulation Conference*, Edited by M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, 427–436. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Kemper, P. 2007. "A Trace-based Visual Inspection Technique To Detect Errors In Simulation Models". In *Proceedings 2007 Winter Simulation Conference*, Edited by S. G. Henderson, B. Biller, M.-H.

- Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, 747–755. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Kleijnen, J. P. C. 1999. “Validation of Models: Statistical Techniques and Data Availability”. In *Proceedings of the 1999 IEEE Winter Simulation Conference*, Edited by P. A. Farrington, H. B. Nembhard, D. T. Sturrock, and G. W. Evans, 647–657. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Mosinski, M., D. Noack, F.S. Pappert, O. Rose, W. Scholl. 2011. “Cluster based analytical method for the lot delivery forecast in semiconductor fab with wide product range”. In *Proceedings of the 2011 Winter Simulation Conference*, Edited by S. Jain, R. R. Creasey, J. Himmelpach, K. P. White, and M. Fu, 1829–1839. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Nazzal, D., A. El-Nashar. 2007. “Survey of research in modeling conveyor-based automated material handling systems in wafer fabs”. In *Proceedings of the 2007 Winter Simulation Conference*, Edited by S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, 1781–1788. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Saanen, Y. 2003. “The Design and Assessment of Next Generation Automated Container Terminals”. In *Proceedings 15th European Simulation Symposium (ESS 2003)*, Edited by A. Verbraeck, and V. Hlupic, Delft, The Netherlands, 577–584.
- Sargent, R. G. 2004. “Validation and Verification of Simulation Models”. In *Proceedings of the 2004 Winter Simulation Conference*, Edited by R. G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, 17–28. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Scholl, W., B. P. Gan, D. Noack, P. Preuss, M. L. Peh, P. Lendermann, O. Rose. 2010. “Towards Realization of a High-Fidelity Simulation Model for Short-term Horizon Forecasting In Wafer Fabrication Facilities.” In *Proceedings of the 2010 Winter Simulation Conference*, Edited by B. Johansson, S. Jain, J. Montoya-Torres, J. Hugan, and E. Yücesan, 2563–2574. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Schwenke C., V. Vasyutynskyy, A. Röder, K. Kabitzsch. 2011. “Analysis of Maintenance Histories of Industrial Equipment with Frequent Maintenance Demand”. In *Proceedings of IEEE 9th International Conference on Industrial Informatics (INDIN 2011)*, Lisbon, Portugal, 299–304.
- Shanthikumar, J. G., S. Ding, and M. T. Zhang. 2007. “Queueing theory for semiconductor manufacturing systems: a survey and open problems”. *IEEE Transactions on Automation Science and Engineering*, vol. 4, no. 4, 513–521.
- Struss P., B. Ertl. 2009. “Post-mortem Diagnosis of Bottling Plants Based on Recorded Data”. In *Safe-process '09, 7th IFAC Symposium on Fault Detection, Supervision and Safety*, Barcelona, Spain.
- van der Aalst, W. M. P, H. A. Reijers, A. J. M. M. Weijters, B. F. van Dongen, A. K. Alves de Medeiros, M. Song, H. M. W. Verbeek. 2007. “Business process mining: An industrial application”. *Information Systems*, vol. 32, no. 5, 713–732.
- Vasyutynskyy, V., A. Gellrich, K. Kabitzsch, D. Wustmann. 2010. “Analysis of Internal Logistic Systems Based on Event Logs”. In *Proceedings of the 15th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA 2010)*, Bilbao, Spain.
- Verbraeck, A., C. Versteegt. 2001. “Logistic control for fully automated large scale freight transport systems; Event based control for the Underground Logistic System Schiphol”. In *Proceedings of the IEEE 4th International Conference on Intelligent Transportation Systems (ITSC 2001)*, Oakland (CA), USA, 768–773.
- Versteegt, C., A. Verbraeck. 2002. “The Extended Use of Simulation in Evaluating Real-Time Control Systems of AGVs and Automated Material Handling Systems”. In *Proceedings of the 2002 Winter Simulation Conference*, Edited by E. Yücesan, C.-H. Chen, J. L. Snowdon, and J. M. Charnes, 1659–1666. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Zhou, Z., O. Rose. 2010. “WIP Balance Study from Viewpoint of Tool Group in a Wafer Fab”. In *Proceedings of the 14th Conference on Simulation in Production and Logistics ASIM*, Karlsruhe, Germany, 341–348.

## **AUTHOR BIOGRAPHIES**

**CLEMENS SCHWENKE** is a PhD student at Dresden University of Technology. He is a member of the scientific staff of Professor Klaus Kabitzsch at the Chair of Technical Information Systems. He received his M.S. degree in Electrical Engineering from Dresden University of Technology. His research interests include automation as well as modeling, analysis and design of discrete event systems. His email address is [clemens.schwenke@tu-dresden.de](mailto:clemens.schwenke@tu-dresden.de).

**THOMAS WAGNER** is a PhD student at Dresden University of Technology. He is a member of the scientific staff of Professor Klaus Kabitzsch at the Chair of Technical Information Systems. He received his M.S. degree in computer science from Dresden University of Technology. His research interests include modeling, analysis and optimization of logistic systems. His email address is [thomas.wagner2tu-dresden.de](mailto:thomas.wagner2tu-dresden.de).

**ANDRÉ GELLRICH** is a PhD student at Dresden University of Technology. He is a member of the scientific staff of Professor Klaus Kabitzsch at the Chair of Technical Information Systems. He received his M.S. degree in computer science from Dresden University of Technology. His research interests include modeling and analysis of logistic systems as well as data storage technologies. His email address is [andre.gellrich@tu-dresden.de](mailto:andre.gellrich@tu-dresden.de).

**KLAUS KABITZSCH** holds the Chair for Technical Information Systems at the Institute of Applied Computer Science of the Dresden University of Technology, Germany. He holds a Diploma and a Ph.D. (1982) in electrical engineering and communications technology. He worked in the following years in industry in the area of control of machines and equipment. After 1989 he worked at the universities of Leipzig and Dresden and focused his research on sensor networks, distributed systems, real time software and automation. He became professor and head of the department of technical computer sciences in 1993 at the Dresden University of Technology. His current projects focus on component based software design, design tools for networked automation, energy and quality management, data analysis, advanced process control and predictive technologies. He is a member of IEEE, VDE and GI. His email address is [klaus.kabitzsch@tu-dresden.de](mailto:klaus.kabitzsch@tu-dresden.de).