

MODELING AND WAFER DEFECT ANALYSIS IN SEMICONDUCTOR AUTOMATED MATERIAL HANDLING SYSTEMS

Thomas Wagner
Clemens Schwenke
Klaus Kabitzsch

Technical University of Dresden
Noethnitzer Strasse 46
D-01187 Dresden, GERMANY

ABSTRACT

Modeling and analysis of automated material handling systems in semiconductor manufacturing is difficult because of its complexity, the large amount of data originating from different sources as well as the often incomplete monitoring of transport processes. This article proposes an automated method and tool for building high level models of such systems based on transport logs, routing information and static system data. On the basis of this model, a method for tracing and correlating lot movements is presented and used to support system experts in locating fab areas that most likely caused defects measured on wafers, e.g. due to temporarily contaminated clean room air. In addition, several methods to analyze the transport systems performance, such as the determination of lot detours or causes for a potentially critical load of certain system parts are discussed.

1 INTRODUCTION

In recent 200mm semiconductor manufacturing, wafer transport processes are executed by a complex automated material handling system (AMHS), which usually consists of several thousand elements. Production lots spend a significant amount of time using these systems (currently up to 20 percent of the overall production time) for transits between production equipments. Because of the growing complexity of semiconductor products (amongst others due to a change in product lines towards customer-oriented, energy efficient chips) and by association a rising amount of production steps, these AMHS will be characterized by higher traffic intensity in the near future, as stated in Nazzal and El-Nashar (2007). In order to reduce the costs of the logistic processes (currently up to 10 percent of overall production costs according to Rushton, Croucher, and Baker 2006), an adequate analysis of the transport processes is necessary to recognize optimization possibilities like bottlenecks or other weak spots. It is also necessary to constantly monitor the system to detect emerging problems like tailbacks or increasing usage of system elements already working at full capacity.

Due to the increasing time spent in the material handling system, there is a growing possibility of wafer defects occurring during the transportation and storage processes (e.g. due to temporarily contaminated clean room air), especially if the lots are transported in open carriers. In case of any anomalies or defects measured on wafers, it is necessary to track the production steps and wafer movements within the relevant time frame (e.g. between two measurement steps). If the defects are not attributable to any single production equipment, it is also necessary to take the exact wafer movement paths within the transport system into account in case the anomalies were caused during those transports. By comparing (correlating) the movements of affected and unaffected lots, the fab (semiconductor manufacturing plant) area most likely causing the measured defects can be narrowed down.

However, the necessary information for such analysis is often incomplete, for example due to insufficient monitoring of lot movements due to the size of the AMHS. Signaling and logging of every movement step of every lot transport (tens of thousands a week using several tens to several hundred transport system elements) may result in a processing and network load no longer manageable by the PLC's (Programmable Logic Controllers) controlling the system. Moreover, a history of several days to weeks of lot movement data would have to be kept, slowing down the analysis due to its size.

In this paper, a flexible method and tool for modeling an AMHS, based on transport system layout and properties, routing information and historical transport data is briefly described. On its basis, an approach to correlate the movement traces of lots to locate fab areas most probably causing measured defects is presented. Furthermore, the suitability of the approach to monitor and analyze critical aspects of logistic systems is discussed and its limits outlined.

The remainder of the paper is structured as follows. Section 2 briefly reviews the related work in the field of the analysis of factory logistic systems, log processing and model building. Section 3 describes the investigated system and shows the overall workflow of the approach. The results are demonstrated in Section 4. Further applications of the developed demonstrator are discussed in Section 5. Finally, a conclusion and outlook are given in Section 6.

2 RELATED WORK

The problem of analyzing complex factory logistic systems was already addressed by the authors, e.g., in Vasyutynskyy et al. (2010) and Wustmann, Schmidt, and Vasyutynskyy (2010). These works focused on the analysis of simulation model runs of these systems based on event logs generated by the simulation. Although the developed methods for automatic model building are still applicable to real factory systems, they rely on the availability of a complete chain of events (resembling the sequence of traversed AMHS elements for each lot) to reconstruct the lot movement paths. In many cases, as in the use case described below, this fine-grained data is not available. The same limitation applies to a paper by Struss and Ertl (2009), describing the post-mortem diagnosis of bottling plants. In addition, the transported goods are not distinguished in this approach, so it is no longer possible to trace single lots.

Other examples or case studies of data mining in trace files (involving the generation of models) can be found in Schwenke et al. (2011) and Röder et al. (2005), although with different scopes. While the former focuses on the investigation of equipment maintenance time intervals (repair duration, fault rates), the latter is dedicated to the analysis of single production equipments.

Further papers in the field of automatic log processing and the analysis of (simulation) traces include Kemper and Tepper (2005), Kemper and Tepper (2009) and Better, Glover, and Laguna (2007). However, these papers focused on the validation of simulation models and not on the analysis of real processes. The analysis of business processes based on event data, for instance discussed in Cook et al. (2004), van der Aalst, Weijters, and Maruster (2004) and van der Aalst et al. (2007), resembles the approach proposed in this paper. Nonetheless, these authors as well as Cook and Wolf (1998), concentrate on processes identification and reconstruction.

The preliminary work described above was used to model the conveyor based material handling system of a 200mm fab. Based on this model, a novel approach to correlate the movement traces of lots to locate fab areas most probably causing measured wafer defects is proposed.

3 DESCRIPTION OF THE APPROACH

The main goal is to identify and visualize wafer defect causes based on defect statistics. A wide variety of corresponding statistical methods and visualizations is readily available within a model-based analysis framework developed by the authors. This framework can be applied to a broad class of logistics systems. Given this precondition, the problem is reduced to mapping proprietary source data to the extended state transition model used within the tool and choosing the algorithms matching the analysis goals.

3.1 Investigated System

The investigated system is a conveyor based continuous flow transport system of a large 200mm semiconductor manufacturing plant. It is subdivided into several production areas (intra-bays connected with inter-bays) and consists of several thousand elements of different types like linear conveyor modules, turntables, elevators or storage elements (stockers). The sequence of transport components to be used by a lot with a certain source and destination (e.g. a transport from a stocker to a productive equipment) is generated by a set of programmable logic controllers (PLC's) controlling the system. In some cases, lots are redirected (transport destination was altered) during a transport. This may be the case if a lot was scheduled for storage in a stocker, but requested from production equipment before they arrive at their original destination. To summarize, the logged events include the lot ID, the starting time and device of a transport job as well as the arrival time and the destination device for each transport. If a redirect event occurred, the corresponding conveyor element and time stamp is also recorded.

In addition, basic transport system layout and property knowledge was provided. Notably, the detailed lot routes during the transports (sequence of traversed AMHS elements) are not monitored. This circumstance prevents an effective lot tracking using only the available transport log data.

3.2 Data Acquisition and Aggregation

The first step to obtain the AMHS model is to consolidate all data in a flexible format suitable for further processing. The analysis framework is based on an extended state transition model because it resembles a least common denominator of different logistic systems, since it complies with the nature of logistic as a sequence of transportation actions. A simplified example is shown in Figure 1.

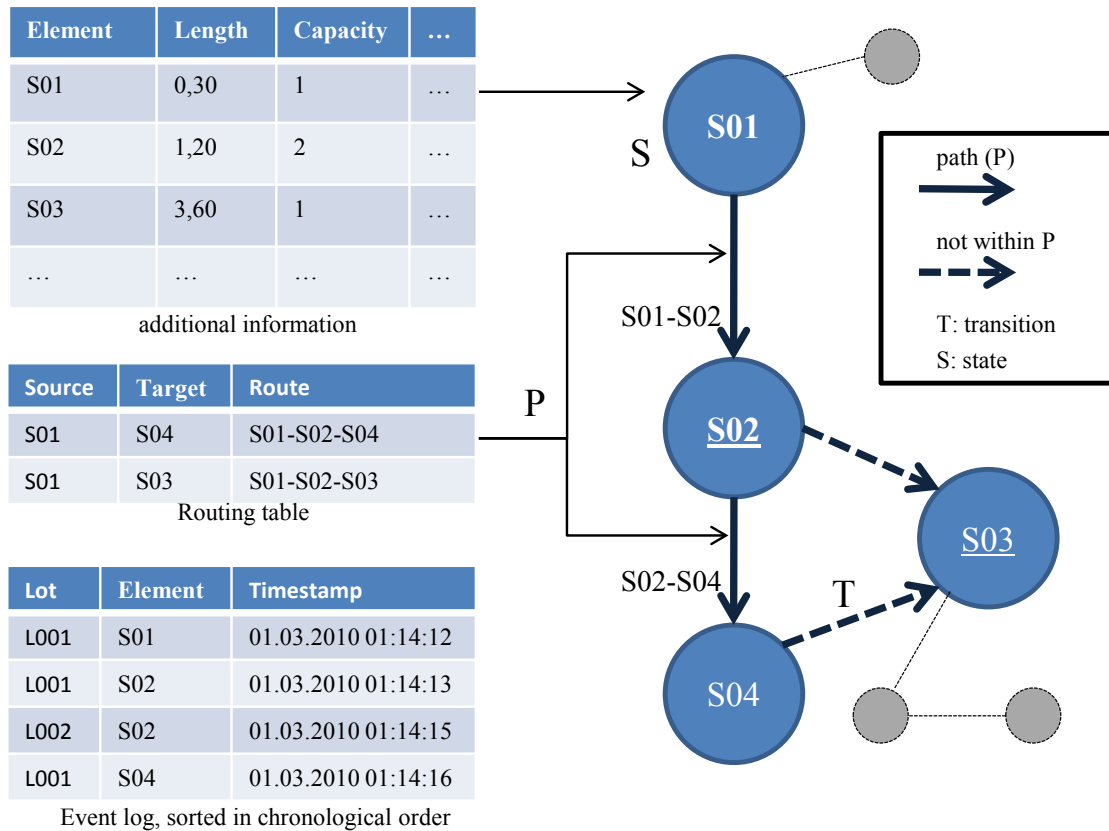


Figure 1: Building the components of the logistic system model.

The set of states S in the graph represents the transport system elements (e.g., turntables or linear conveyor modules) and the set of transitions T resemble the connections between the single elements (a turntable can be used as a crossing or to unite or split transport stream, therefore it is connected to several other elements and owns multiple transitions). In many scenarios, this model can be learned from historical log data. However, in the use case at hand this is not possible because the transport log does not contain enough information (a complete chain of events resembling the sequence of traversed AMHS elements for each lot would be needed). Instead of reconstructing the sequence of states from the event data, a software module transforming the available routing algorithm was used to construct T and the set of paths P (collection of transport routes used by the lots).

For example (in case of a lot being redirected en route), a path from the start state to the state the redirect was conducted is associated to the start event and a path from the redirect to the target state is associated to the redirect event. This method circumvents the need for storing an event for every movement step, thus saving storage space and query effort. An excerpt of the resulting model, roughly resembling the original material handling system, is shown in Figure 2.

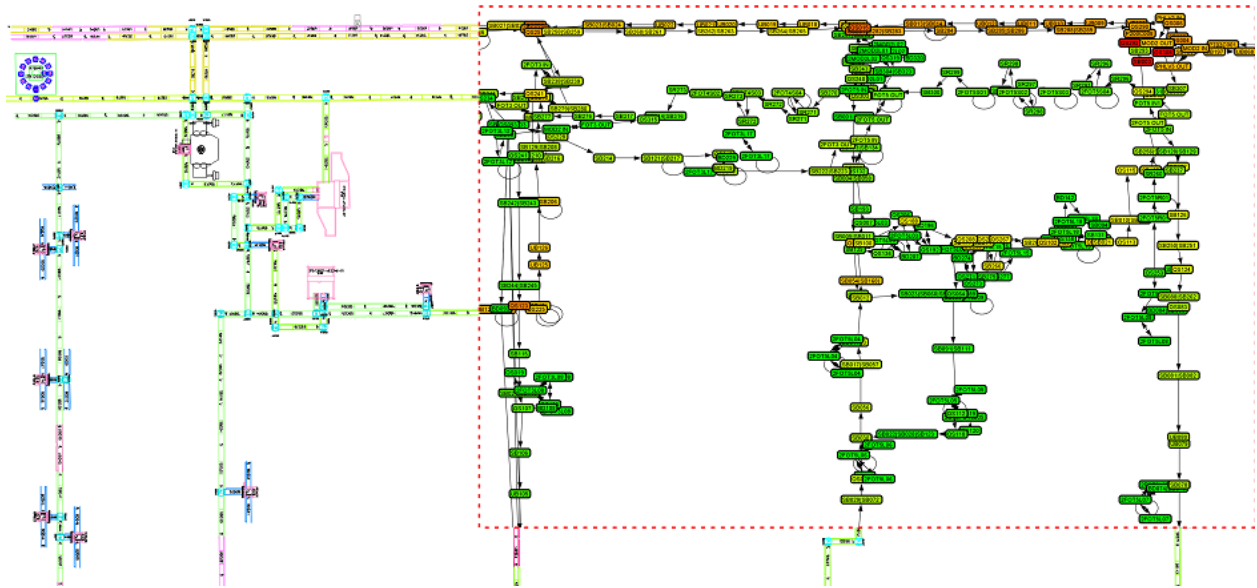


Figure 2: Excerpt of the generated model.

3.3 Visualization and Correlation of Lot Movements

Using the acquired data it is now possible to visualize selected lot movements and their starts and destinations, e.g., the movements of affected (or unaffected) lots in a given time frame, to get a first indication of areas of the fab the problems may (or may not respectively) have occurred. If transport system elements are used by all affected but not by any unaffected lots (within the critical timeframe), the location of the problem source will most likely be near those components, provided that the defects have indeed been introduced during wafer transports. The probability should decrease with an increasing amount of unaffected lots as well as a decreasing amount of affected lots traversing them in the same time span. Figure 3 depicts a simplified example.

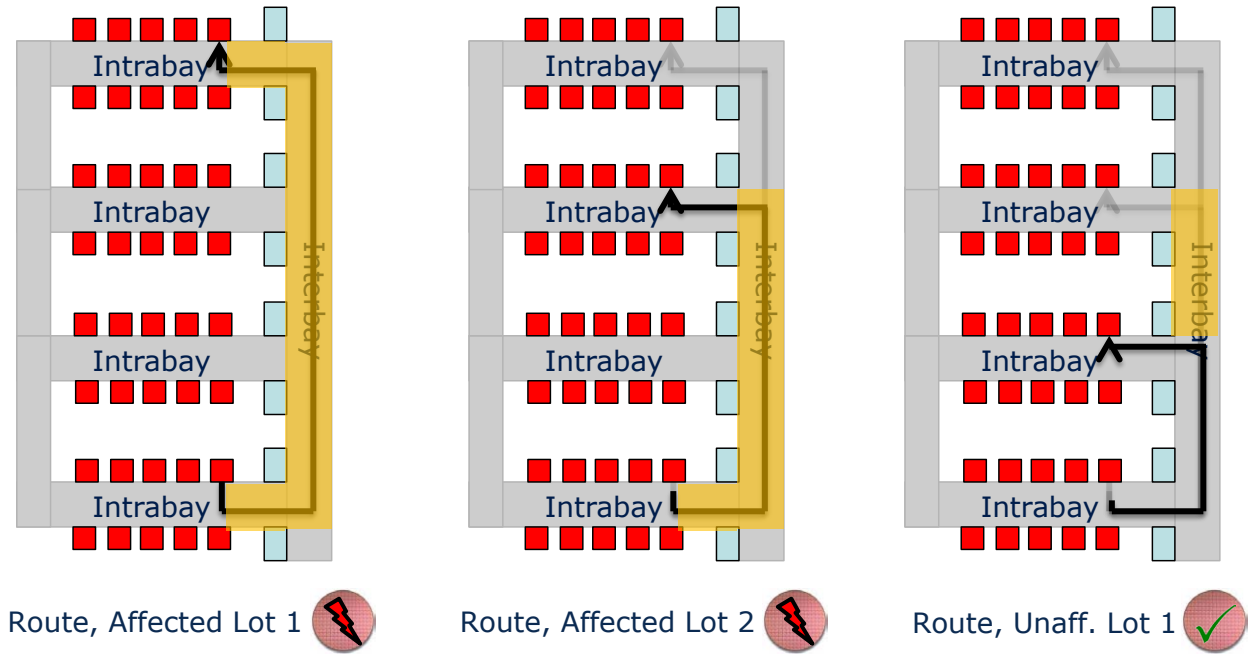


Figure 3: Localizing possible defect-causing areas (orange rectangle) within a transport system by the use of transport routes.

In this example, the area most probably causing the wafer defect (orange rectangle) is at first assumed to be in the vicinity of parts of the conveyor system, that were used for transports of affected Lot Number one in the relevant time span (Figure 3, a). If other affected lots are known, the area can be further narrowed down to the location of conveyor elements commonly used by all of them (Figure 3, b). Lastly, the transports and routes of unaffected lots can be used to further reduce the amount of conspicuous system parts, which likely would have caused the same defect otherwise (Figure 3, c).

Assuming the availability of information about both affected and unaffected lots and critical time spans, one approach to further support the system expert (in identifying suspicious fab areas) can therefore be the correlation of transport element usage statistics with the knowledge of a lot being affected or unaffected by an anomaly.

3.4 Correlation Approach

To compute the correlation between the transport element usage and the wafer defect affection (i.e., if a wafer or lot was affected by the investigated defect), a routing matrix and an affection vector were constructed. The following steps are required to produce them:

1. The routing matrix is initialized to contain a column for each conveyor element present in the system.
2. All transport events (Start, Stop, Redirect) and the corresponding routes from lots are fetched, for which information on their affection state is available within the critical time span.
3. For each relevant lot, a new line is created in the routing matrix. Each cell is filled with either the value one if at least one of the routes of this lot contains the appropriate conveyor element (the lot traversed it during this particular time span) or zero otherwise.
4. For each relevant lot, the value one is added to the affection vector if the corresponding lot is defective or zero otherwise.

5. After all routes are processed, columns containing only zero values are deleted from the routing matrix because no affected lots used them in the relevant time frame, practically excluding them from the list of suspects.

Each of the columns of the resulting matrix was then correlated with the affection vector to obtain a correlation coefficient for each transport system element, denoting its association to the affection state of the lots. Since the two variables (conveyor element usage- and affection vector) are both binary, the authors decided to use the phi coefficient ($\hat{r}_\phi$) as a special case of the Pearson Product Moment Correlation, if used with two dichotomous variables (Cramér 1999). The necessary contingency tables, containing the frequency distributions of the features, are built as shown in Table 1.

Table 1: The contingency tables required to calculate $\hat{r}_\phi$, where a is the sum of unaffected lots that did not traverse the corresponding element, b is the sum of affected lots that traversed the element and so forth.

	Lot not affected	Lot affected
Element not traversed	a	b
Element traversed	c	d

The phi coefficient can then be calculated according to (1) for each conveyor element. For a simple example, the procedure is illustrated in Figure 4.

$$\hat{r}_\phi = \frac{(a \cdot d) - (b \cdot c)}{\sqrt{(a+b) \cdot (a+c) \cdot (b+d) \cdot (c+d)}} \quad (1)$$

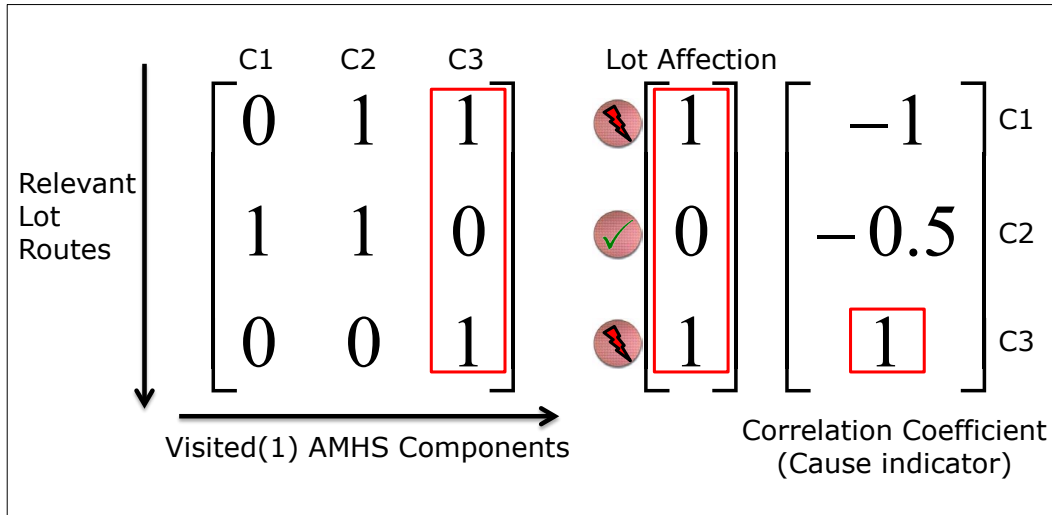


Figure 4: Simplified example of the correlation procedure.

The resulting values for each transport system element can then be interpreted as the likelihood, that a particular transport system element caused the measured defect as follows:

- $-1 \leq \hat{r}_\phi \leq 0$: The particular element was mostly used by unaffected lots, low probability
- $0 \leq \hat{r}_\phi \leq 1$: The particular element was mostly used by affected lots, increasing probability

Each transport system element was then colored accordingly from the highest (red) to the lowest (green) observed $\hat{r}_\phi$ and visualized within the transport system model, providing visual hints about fab areas

possibly causing the measured defects. Figure 5 shows an example screen shot taken during an ongoing investigation of a measured wafer defect.

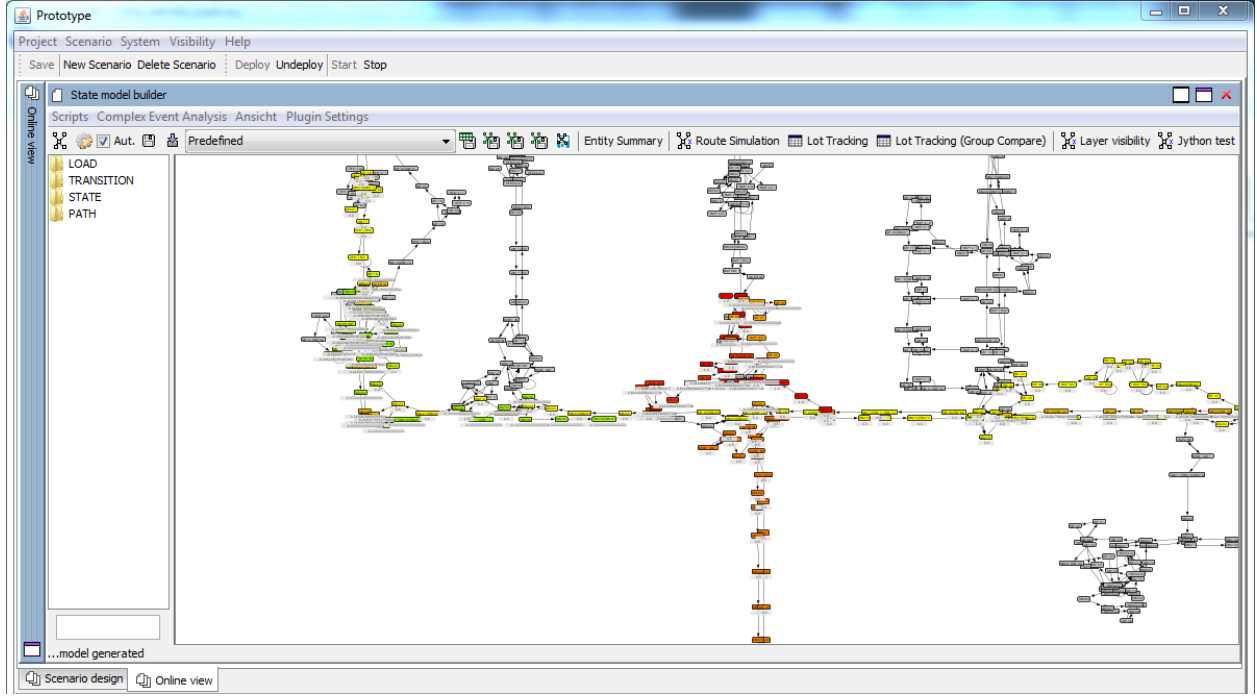


Figure 5: Visualization of the wafer defect indicator (red: transport system elements most probably causing a measured defect) within the AMHS (excerpt).

4 APPLICATION AND RESULTS

The calculated lot movements and paths have been validated by comparing them with measurements of track readers placed in selected parts of the material handling system. The effectiveness of the correlation approach is tested during ongoing and future defect investigation efforts.

To test the effectiveness of the approach without having to wait until a defect occurs, historical transport data from the fab has been used to simulate defects. To reduce the amount of computation time needed, a subset of the transported lots and the corresponding routes was randomly chosen from the history of a 24 hour time slot. If such a lot passed a particular transport system element, it gets marked as “defective” with a certain probability. The transport system elements (causing the defects) were selected randomly from the classes of highly (e.g. the inter-bay) and seldom (e.g. the intra-bay) used parts of the system. The algorithm has then been applied to a range of test cases and the results of tree experiments are shown in Table 2.

Since the correlation coefficients differ from test run to test run (the lots are randomly chosen each time), the rated results are presented instead. For this purpose, the transport system elements were assigned to transport system areas. An area is a section of the transport system between two junctions, e.g. a strain of linear conveyor elements between two turntables, since the correlation algorithm is not able to distinguish those elements (i.e., the same value of $\hat{r}_\phi$ is assigned to all of them) because of the missing routing diversity. These areas are listed according to the resulting correlation coefficient in descending order (i.e., the most suspicious areas are at the top of this list). The *Rating* value is the position of the simulated defect causing area in this list.

Table 2: Test results after applying the algorithm to different use-cases.

Lots investigated		Rating of the defect causing area (see below)		
		150 Lots	60 Lots	12 Lots
Defect probability (percent)	40	1	1	1
	35	1	1	1
	30	1	1	2
	25	1	2	>30
	20	1	1	>30
	15	1	3	>30
	10	5	9	>30
	5	11	17	>30
	1	26	30	>30

The results show that a certain amount of affected and unaffected lots is required to produce a satisfactory solution. The results for the experiments with twelve lots can be considered unhelpful, although they implicate very good results with the knowledge of only two to three defect lots, because the same value of $\hat{r}_\phi$ is assigned to very large areas of the transport system (i.e., the defect causing area cannot be narrowed down sufficiently). This can be explained by the lack of diversity in the movement paths of only three lots. In the use case at hand, the knowledge of around seven to eight affected lots and the same amount of unaffected lots proved to be sufficient.

In addition, there is a possibility of “false negatives” (conveyor elements erroneously showing a low degree of correlation to the affection state of lots) being produced. This might happen in case the time span in which the defect occurred cannot be narrowed down satisfactorily, e.g. due to unusually long intervals between measurement steps. In an example scenario, most of the unaffected lots could have used a particular conveyor element before the critical event (causing the defect) occurred. In that case, the correlation coefficient calculated would be lower than expected, even though all affected lots might have traversed this element after the event. Such a use-case has been simulated as shown in Table 3.

Table 3: Test results after applying the algorithm to different use-cases.

		60 Lots, 30% Defect Probability
		Rating of the defect causing area
Defect lasted (from 24h)	hours	
	24	1
	20	1
	16	1
	12	4
	8	10
	4	30

In this example, the approach was not able to correctly identify the defect causing area if the defect lasted only about half the time of the investigated time span. This issue could be addressed by applying the proposed algorithm on time slices obtained by subdividing the relevant time span into smaller parts once the need arises.

Although the approach produced convincing results so far, there are some constraints to be considered before applying the approach. To summarize:

- The investigated transport system network has to be interconnected to a certain degree.
- The more the transport routes of the investigated lots differ, the better the achievable results are.

- A certain amount of affected and unaffected lots is required to produce a satisfactory solution, the necessary amount of lots depends on the size of the system and the diversity of the transport routes.
- The time span in which the defect occurred should be narrowed down sufficiently to improve the results.

5 FURTHER APPLICATIONS

In the use case stated in Section 3.1, additional information about the logistic system (such as conveyor element positions, capacities and lengths) was available. Therefore, further analysis algorithms could be applied without changing the underlying model, for example:

- Measuring the degree of capacity utilization of each conveyor element, testing for upward trends in the usage statistics of critical (highly used) elements.
- Checking for major overlaps (a lot of commonly used segments leading to bottlenecks) in the observed wafer routes, which could be prevented by adjusting the routing algorithms or by extending the material handling system.
- Detecting lot detours by comparing equipment distances with the lengths of the ascertained paths.

These analysis methods are used to monitor the performance of the material handling system and to generate suggestions for further optimization. Figure 6 shows some of the results.

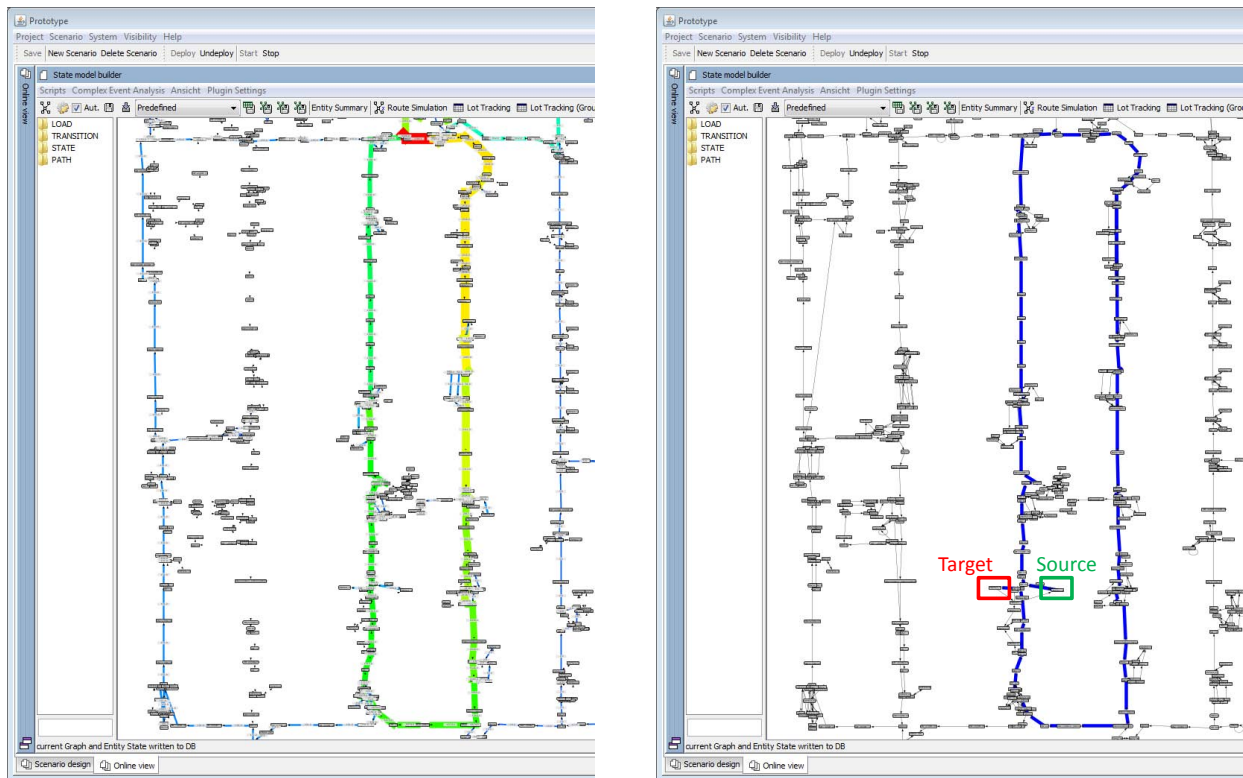


Figure 6: Visualization examples of analysis methods (left picture: lot traffic streams within a time span as gradient colors, right picture: visualization of a lot detour path as a blue line).

Although the analysis capabilities of the tool can be extended and automated with the use of scripts and rules, the possibilities are somewhat limited due to the nature of the data. Since only two to three events (transport start, stop and redirect, if applicable) are available for each transport, dynamic effects

(e.g., tailbacks, transient delays or route deviations due to local congestion avoidance strategies on PLC level) are not observable. A reconstruction of these effects, for example to predict future overload situations for the purpose of supporting dynamic routing decisions, could be possible using a detailed simulation of the logistic system. However, it would be very complicated to create and maintain a simulation of such a complex system (Kohn et al. 2009). To enable the detection and analysis of such effects, it is planned to refine the granularity of the events reported by the PLC's.

6 CONCLUSION AND OUTLOOK

Up to now, the ascertainment of fab areas possibly causing wafer defects during lot transports, even in comparably small periods of time, is a tedious and time consuming task in many applications where appropriate monitoring data, such as complete wafer tracking or a thorough monitoring for clean room air contamination, is unavailable. The presented approach enables system experts to easily reproduce and visualize lot movements in a timely manner. It supports them in solving potentially production critical situations by outlining areas of the transport system which have to be examined further. Because of the underlying abstract modeling method (state-transition model), the tool and the method described can easily be adapted to other logistic systems, provided that information about transport jobs, routing data and information about relevant lots (e.g., damaged and undamaged goods) is available. By using other available information about the logistic system at hand, several performance analysis algorithms could be applied (e.g., the detection of heavily used parts of the transport system or wafer detours within the system). These studies will serve as a base for planning transport system improvements in the near future.

Although it was possible to model the material handling system to an adequate degree by using only a small set of input data (thus avoiding efforts and costs of refining the monitoring of transports by subsequent installation of new hardware components) it is evident that the granularity of the available events is not sufficient to perform analysis relying on the knowledge about dynamic effects. Therefore, the next tasks will include:

- Refining the correlation approach to better cope with cases where the critical time span (in which the wafer defect occurred) cannot be narrowed down.
- Investigate possibilities to increase the granularity of the AMHS monitoring (i.e., the amount and quality of the events produced) without overloading the controllers.
- Using the refined monitoring data to apply methods for detecting dynamic effects (e.g. delays), which are readily available within the framework.

The knowledge about dynamic effects in the conveyor system could then be used e.g. to notify the system expert about arising problems (prediction of tailbacks) or serve as an input of dynamic routing algorithms, allowing them to bypass problematic conveyor elements.

ACKNOWLEDGEMENTS

This work originates in the context of the research and development project Cool Fab (Grant Nr. 13N10770) funded by the German Federal Ministry of Education and Research (BMBF) within the Leading-Edge Cluster Cool Silicon.

REFERENCES

- Better, M., F. Glover, and M. Laguna. 2007. "Advances in analytics: Integrating dynamic data mining with simulation optimization". *IBM Journal of Research and Development* 51 (3.4): 477–487.
- Cook, J. E., Z. Du, C. Liu, and A. L. Wolf. 2004. "Discovering models of behavior for concurrent workflows". *Computers in Industry* 53 (3): 297–319.

- Cook, J. E., and A. L. Wolf. 1998. "Discovering Models of Software Processes from Event-Based Data". *ACM Transactions on Software Engineering and Methodology* 7:215–249.
- Cramér, H. 1999. *Mathematical methods of statistics*. Princeton landmarks in mathematics and physics. Princeton University Press.
- Kemper, P., and C. Tepper. 2005, December. "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., and C. Tepper. 2009. "Automated Trace Analysis of Discrete-Event System Models". Volume 35, 195–208. Piscataway, NJ, USA: IEEE Press.
- Kohn, R., D. Noack, M. Mosinski, Z. Zhugen, and O. Rose. 2009, December. "Evaluation of modeling, simulation and optimization approaches for work flow management in semiconductor manufacturing". In *Proceedings of the 2009 Winter Simulation Conference*, edited by M. D. Rossetti, R. R. Hill, B. Johansson, A. Dunkin, and R. G. Ingalls, 1592–1600. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Nazzal, D., and A. El-Nashar. 2007, December. "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.
- Röder, A., V. Vasyutynskyy, K. Kabitzsch, T. Zarbock, and G. Luhn. 2005. "Log-based State Machine Construction for Analyzing Internal Logistics of Semiconductor Equipment". In *Proceedings of the International Conference on Modeling and Analysis of Semiconductor Manufacturing*, 54–60.
- Rushton, Croucher, and Baker. 2006. *The handbook of logistics and distribution management*. Kogan Page Series. Kogan Page.
- Schwenke, C., V. Vasyutynskyy, A. Röder, and K. Kabitzsch. 2011. "Analysis of maintenance histories of industrial equipment with frequent maintenance demand". In *Proceedings of the 9th IEEE International Conference on Industrial Informatics (INDIN 2011)*, 299–304.
- Struss, P., and B. Ertl. 2009. "Post-mortem Diagnosis of Bottling Plants Based on Recorded Data". In *Safeprocess '09, 7th IFAC Symposium on Fault Detection, Supervision and Safety*. Barcelona.
- van der Aalst, W. M. P., H. A. Reijers, A. J. M. M. Weijters, B. F. van Dongen, A. K. A. de Medeiros, M. Song, and H. M. W. Verbeek. 2007. "Business process mining: An industrial application". *Information Systems* 32 (5): 713–732.
- van der Aalst, W. M. P., T. Weijters, and L. Maruster. 2004. "Workflow mining: discovering process models from event logs". *IEEE Transactions on Knowledge and Data Engineering* 16 (9): 1128–1142.
- Vasyutynskyy, V., A. Gellrich, K. Kabitzsch, and D. Wustmann. 2010. "Analysis of internal logistic systems based on event logs". In *Proceedings of the 15th IEEE Conference on Emerging Technologies and Factory Automation (ETFA 2010)*.
- Wustmann, D., T. Schmidt, and V. Vasyutynskyy. 2010. "Identification and quantification of anomalies in the behaviour of material flow systems based on event logs". In *Proceedings of the 2010 ASIM symposium "Simulation in Produktion und Logistik"*, 429–436. Karlsruhe: KIT Scientific Publishing.

AUTHOR BIOGRAPHIES

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 e-mail is thomas.wagner2@tu-dresden.de.

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 at the

Department of Computer Science. 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 e-mail is clemens.schwenke@tu-dresden.de.

KLAUS KABITZSCH is the head of the department of technical computer sciences at the Dresden University of Technology. He received a Diploma and a Ph.D. in electrical Engineering and communications technology. His current projects have their focus on the automation domain, 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 e-mail is klaus.kabitzsch@tu-dresden.de.