Simulation in Production
and Logistics 2015
Markus Rabe & Uwe Clausen (eds.)
Fraunhofer IRB Verlag, Stuttgart 2015

Avoiding Equidistances when Routing by Shortest Paths

Vermeidung von Äquidistanzen beim Routing nach kürzesten Wegen

Christian Hammel, Robert Schmaler, Thorsten Schmidt, Technische Universität Dresden, Dresden (Germany), christian.hammel@tu-dresden.de, robert.schmaler@tu-dresden.de, thorsten.schmidt@tu-dresden.de

Matthias Schöps, Globalfoundries Module One LLC & Co. KG, Dresden (Germany), matthias.schoeps@globalfoundries.com

Abstract: A considerable part of automated material handling systems (AMHS) use some sort of shortest path routing for defining the transport path. If the shortest paths are ambiguous (multiple shortest paths for one relation) there is often no explicit rule. Therefore, the exact behaviour is not completely predictable in simulation models. This reduces usability of simulation for analysing and optimising system behaviour, especially when examining behaviour on over-utilised tracks. To assess the potential impact of multiple shortest paths, a method is presented identifying all relevant node pairs. Also a method to eliminate these equidistances is discussed. Both methods are tested in AMHS simulations of an existing semiconductor fab. The case study also covers an impact analysis of multiple shortest paths in a real system. This shows outcome deviations of more than 10% in terms of throughput or delivery times – solely dependent on the exact behaviour in case of ambiguous paths.

1 Introduction

In areas like the semiconductor industry requirements to be fulfilled by installed material handling systems are steadily rising. This results in bigger and more complex systems. Current automated material handling systems (AMHS) in modern semiconductor fabs comprise a track system of several kilometres in length, more than 1,000 intersections and use several hundred vehicles to carry out some thousand transport jobs per hour (Gaxiola et al. 2012; Hammel et al. 2012).

The immense investment in the tool outfit in semiconductor industry makes it especially important to avoid any inefficiency in the production process. This includes the transportation of material between each two of more than 1,000 process steps. Because of the complexity simulation models are indispensable when checking and optimising the material handling system's functionality and performance. The routing strategy and its parameters are a key component of the transport system to be

reproduced in the model. To provide reasonable results in simulation runs this means the vehicles have to choose the same route in the model as they do in the real system. The exact reproduction may be challenging if the routing is based on shortest paths and these shortest paths are not unique. Differences in throughput and delivery times of more than 10 % have been revealed in simulation runs of a real world use case (see chapter 4). These differences just arose from differing routing schemes that resulted from exactly the same cost parameters, solely the selection from the equidistant paths varied. Hence, usability of simulation as prediction instrument for system performance in such cases suffers considerably.

This paper will introduce the routing problem in material handling systems in chapter 2. In chapter 3 the problem of equidistant shortest paths is described along with an algorithm to identify all origin-destination relations with multiple shortest paths and an approach to eliminate them. The method is demonstrated for an AMHS of a running semiconductor fab in chapter 4. The paper ends with concluding remarks in chapter 5.

2 Routing in Automated Material Handling Systems

In many existing automated material handling systems routing is based on some sort of static shortest path approach. This means all path segments are associated with a parameter based on e.g. length, travel time or some combination. This parameter is understood as costs for using this segment. Algorithms are then used to identify the path from origin to destination for which the total cost is minimal. Hence, 'shortest' path explicitly means the cheapest path regarding these cost parameters, not the one with physically shortest length.

Even though the system size would suggest more sophisticated (dynamic) approaches, necessary communication effort and the existing IT infrastructure mostly prohibit this. Sometimes simple functions are implemented to circumvent emerged traffic congestions, hence only acting reactively. More elaborate countermeasures require dynamic routing approaches (see e.g. Bartlett et al. 2014).

The mentioned shortest path routing is rather simple to implement in simulation models – as long as the shortest paths for all origin-destination relations are unique. If multiple different shortest paths (by cost parameters) exist additional information is needed to decide which of these paths to select. However, in many cases this additional information is not available. In real systems the existence of equidistant shortest paths often seems to be ignored, i.e. no explicit function covers this case. The decision of which of the shortest paths to choose is rather made implicitly. It depends on the specific implementation of the used search algorithm and the storage or search order of the path segments. As this is rarely documented to the last detail it is not possible to reproduce it in simulation models. This problem cannot be solved in the simulation alone. A possibility to avoid this problem right from the start is to eliminate equidistant paths in the real system. At the same time this would provide a higher predictability of system behaviour but it also means affecting the utilisation of path segments as some of the cost parameters have to be adjusted slightly. Especially if the utilisation of path segments is to be optimised by manipulating the cost parameters (Hammel et al. 2012) any such influence has to be examined carefully.

For analysing routing in general and shortest paths in particular a graph model of the transport network is an appropriate measure. Therefore, a graph $G(V,E)$ with vertices V and edges E is constructed by modelling all intersections as well as in- and outputs as vertices $v_i \in V$ and the interconnecting paths / pieces of track by edges $e_j \in E$. Each of the edges is furthermore associated with a cost parameter $c_j = c(e_j)$. A path p_{odk} is a sequence of edges from origin vertex v_o to destination vertex v_d with path costs $c(p_{odk}) = \sum_{e_j \in p_{odk}} c_j$. The distance d_{od} between two vertices v_o and v_d is defined as the path costs of the shortest path, i.e. the path with minimum path costs $d_{od} = \min_k(c(p_{odk}))$. Accordingly, the set of shortest paths from v_o to v_d is given by $MP_{od} = \{p_{odk} | c(p_{odk}) = d_{od}\}$. If MP_{od} consists of at least two paths they are called equidistant paths for relation (v_o, v_d).

3 Equidistant Shortest Paths

In material handling systems experience of the authors tells that the existence of multiple shortest paths for single origin-destination relations is mostly ignored. This may be (if the case is anticipated anyway) because occurrence of such equidistances is assumed scarcely and only little impact to global system behaviour and performance is expected. Furthermore, in the original idea of routing by shortest paths it does not make a difference which of the shortest paths to take, the distance is the same. However, this becomes an issue in case congestions impact travel times. E.g. uncertainty in terms of routes selected by vehicles may cause problems when traffic capacity of single pieces of track gets relevant. A real world application described in chapter 4 shows that the impact of equidistances may in fact be substantial. Additionally, algorithms often used for finding the shortest path (Dijkstra 1959; Floyd 1962; Warshall 1962) are designed to identify a shortest path; to be exactly i.e. one of the potentially multiple shortest paths. In their original form they are explicitly not checking for equidistances. Hence, computer implementations neither throw errors in such cases.

3.1 Issues for Simulation Studies

As mentioned above simulation studies, in particular discrete event simulations, are indispensable for some AMHSs. They are employed e.g. to predict maximum throughput, delivery times and impacts of different control strategies/functions. Even though simulation models as all models are abstractions of real world systems and the level of abstraction may vary, correct reproduction of routing decisions is a basic requirement for the reliability of such predictions.

In case of existing equidistances when routing happens according to shortest paths the accuracy of simulation studies to predict real system behaviour and subsequently the value of such studies is questionable. For assessment it is important to check for equidistances first. Even though this problem affects simulation and its value it is hard to solve without changes in the real system. Achieving this would require implementing the same procedures in the simulation model as occurring in the real system in case of more than one shortest path. While the identification of shortest paths in case they are unique is independent of which search algorithm is used, for multiple shortest paths the identified one might differ according to different algorithms used. Even for one and the same algorithm the respective implementation

and storage of the network structure might give different results. E.g. Dijkstra's algorithm chooses the first shortest path it finds. This means it depends among others on the order it searches through the edges. Consequently, reproducing real system routing in simulation is nearly infeasible in this case.

3.2 Identification

In order to describe the occurrence of multiple shortest paths in a network graph it is sufficient to list a minimum set M of origin-destination relations in which for each node pair the following holds true: origin and destination nodes are the only nodes appearing in all shortest paths. Given that this minimum set of node pairs and their sets of shortest paths are known, all shortest paths for any other node pair can be generated from any one of the shortest paths by substituting respective parts. In particular this means only diverts appear as origins and merges as destinations in this minimum set. In Fig. 1 M would consist of node pairs (A,D), (E,H), (E,J) and (G,J). Even though e.g. for (A,J) the shortest path is not unique either, all shortest paths can be generated if the mentioned set is known. Moreover, an empty minimum set means that there are no equidistances in the network in general.

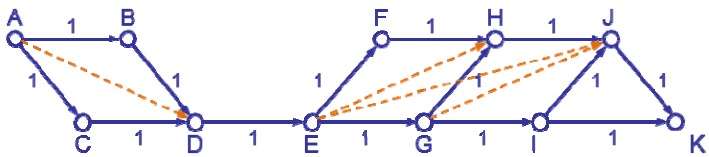

Figure 1: *Graph with uniform edge costs, the dashed lines depict those node pairs of the minimum set of node pairs with multiple shortest paths sufficient to describe all equidistances in the graph*

For finding this minimum set M of node pairs a two-step method is proposed: The first step is a modification of the Floyd-Warshall algorithm commonly used to generate a matrix consisting of the next node on a shortest path for all node-to-node relations. Out of this a shortest path can be reconstructed. Storing not one but multiple next nodes in one position of the matrix in case of multiple shortest paths (if new distance is not less but equal to old distance) allows reconstruction of all shortest paths (Alg. 1).

Algorithm 1: *Modified Floyd-Warshall algorithm for finding all equidistant shortest paths*

```
matrix D with d_ij = minimum cost of edge from v_i to v_j,
                     ∞ if no edge between v_i to v_j
matrix N with n_ij = {v_j} if d_ij < ∞, NULL else
for i from 1 to |V|
     for j from 1 to |V|
          for k from 1 to |V|
               if d_ji + d_ik < d_jk then
                    n_jk = n_ji
                    d_jk = d_ji + d_ik
               else if d_ji + d_ik = d_jk then
                    n_jk = n_jk ∪ n_ji
```

The entries n_{ij} of N provide all next nodes on any shortest path from v_i to v_j. Hence, in the second step only those node pairs for which n_{ij} consists of multiple nodes have to be considered when searching for the minimum set M of pairs describing all equidistances. Constructing all shortest paths from v_i in the direction of v_j by ascending length (similar to the approach used in the Dijkstra algorithm) and checking for the first node v_k where all shortest paths combine again gives a node pair (v_i, v_k) to include into M (Alg. 2).

Algorithm 2: *Finding the elements of the minimum set M of all relevant node pairs for which the shortest path is not unique (run time optimisations omitted for better understanding)*

```
set M := {}
for i from 1 to |V|
     for j from 1 to |V|
          if |n_ij| > 1 then
               set LV := n_ij (ordered list)
               set LD := () (ordered list)
               for k from 1 to |LV|
                    v_s := LV_k
                    append d_is to LD
               while |LV| > 1
                    v_s := LV_k with LD_k = min(LD)
                    for r from 1 to |n_sj|
                         v_t := n_sj,r
                         if v_t ∉ LV
                              append v_t to LV
                              append d_it to LD
                    delete LV_k from LV
                    delete LD_k from LD
               v_s := LV_1
               if (v_i,v_s) ∉ M
                    add (v_i,v_s) to M
```

3.3 Avoidance Approaches

Avoiding/eliminating equidistances means adjusting cost parameters in order to make all shortest paths unique ('shortest' regarding the path with minimum costs). Cost parameter settings can theoretically be constructed in a way that all shortest paths are unique. If e.g. the edges e_j are ordered from $j = 1..n$ and their costs are defined by $c(e_j) = 2^j$ the cost of a path is determined by the edge with the highest index as $2^j > \sum_{i=1}^{j-1} 2^i$. However, in applications such an approach is infeasible because of limited ranges of used data types.

In real world systems the approach for eliminating equidistances from an existing shortest path routing scheme depends on the minimum steps when changing cost parameters. If this minimum step is (arbitrarily) small cost parameters may be adjusted to eliminate equidistant shortest paths without affecting the general routing. This can be achieved by increasing cost parameters for edges associated with all but one of the shortest path for each relevant node pair by an amount small enough not to raise the path costs to or above the second shortest paths (one example shown in Fig. 2). In case of a minimum step (e.g. integer values with a certain upper limit) it is more complicated to avoid impact to the general routing. Potentially any cost increase of one edge could change the shortest paths of all origin-destination pairs previously using this edge, i.e. all second-shortest paths may then have the same costs. This effect could be lessened by multiplying all cost parameters by an integer value first (e.g., by 10) thus generating gaps between the costs of paths and decreasing the probability of impact on the general routing. But the allowed range for path costs has to be considered before.

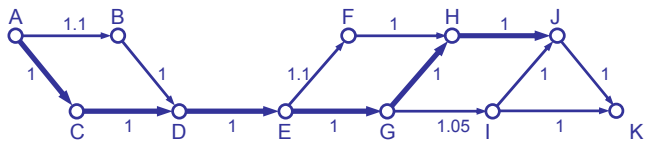

Figure 2: *Example for cost parameter adjustments to eliminate multiple shortest paths by small amounts with shortest path for* (A, J) *in bold*

If impact on the general routing cannot be prevented completely, avoiding equidistances must be incorporated in the method generating the cost parameters in the first place. One possibility is the alternating application of the optimisation method and an iterative approach for eliminating equidistances until the requirements of both parts are fulfilled. The latter iteration would first generate the minimum set of relevant node pairs M and then increase all but one outgoing edge of shortest paths starting at all origin nodes in M (or all but edge of shortest paths related to the incoming edges of all destination nodes in M). For the graph in Fig. 1 for example cost increases in one of the outgoing edges of nodes A, E and G would be sufficient to eliminate all equidistances. These steps are repeated until the generation of M provides an empty set. Fig. 3 shows the iteration for the example graph of Fig. 1. It also illustrates the potential of changing the original routing by the adjustments. The bottom part shows the new shortest path for (G, K) which originally ran through node I only.

As the identification of a solution cannot be guaranteed an appropriate exit condition should be implemented. For a better compatibility of the optimisation and equidistance elimination approaches the optimisation could e.g. increase cost parameters in multiples of the minimum increment, e.g. in even numbers only.

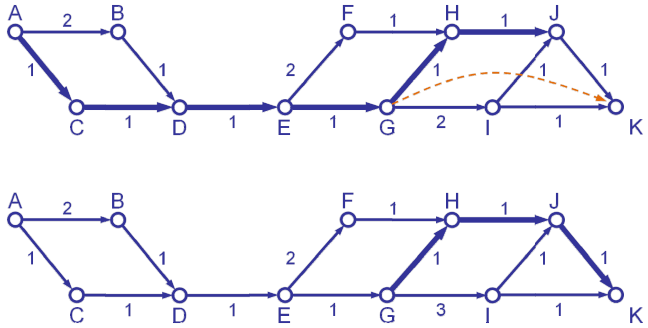

Figure 3: *Example for cost parameter adjustments to eliminate multiple shortest paths by minimum increments (integer values) generating new multiple shortest paths for* (G, K) *in the first step (top, with shortest path for* (A, J) *in bold) and changing the shortest path for* (G, K) *(bold) in the second step (bottom)*

4 Case Study

The described facts and methods will be demonstrated on one of GLOBALFOUNDRIES' overhead hoisting transport systems in a fabrication plant in Dresden, Germany. It has been installed in 2004 and covers the whole production area of one building. It consists of 6,500 metres of track, nearly 1,000 intersections and 280 vehicles. Currently it performs around half a million transport jobs a week.

Vehicles in this system select the specific route to execute an assigned job by employing the Dijkstra algorithm. Cost parameters are therefore assigned to each piece of track. In the original setting these parameters roughly resembled the length of the respective piece of track with some manual adjustments, for computational reasons integer values were set analogue to full decimetres. This rounding is one of the factors raising the probability of equidistant paths occurring. Another factor is the mostly parallel arrangement of tracks in the same direction with several shunts to change from one to the other lane (Fig. 4).

The AMHS has been installed more than ten years ago and requirements especially in terms of throughput have risen (and still further rise) because of new or upgraded tools. Despite some technical upgrades also to the AMHS optimisation measures have to be taken in the control domain as well to be able to fully support production needs. Therefore, an optimisation of track utilisation by adjusting cost parameters as described by Hammel et. al (2012) should be executed. In the optimisation described in the paper equidistances have not been a big issue because for the highly utilised tracks cost parameters had been changed several times before so the occurrence was limited to areas where their impact was indeed negligible. In this new optimisation run cost parameters were to be reset to track lengths before adjustment. After this

reset and the subsequent optimisation the equidistance identification described in section 3.2 showed 162 intersection pairs with multiple shortest paths (Fig. 4). Especially for the typically highly utilised tracks connecting different areas (like in the far right in the figure) these equidistances may have unpredictable consequences as track utilisations are close to their capacity.

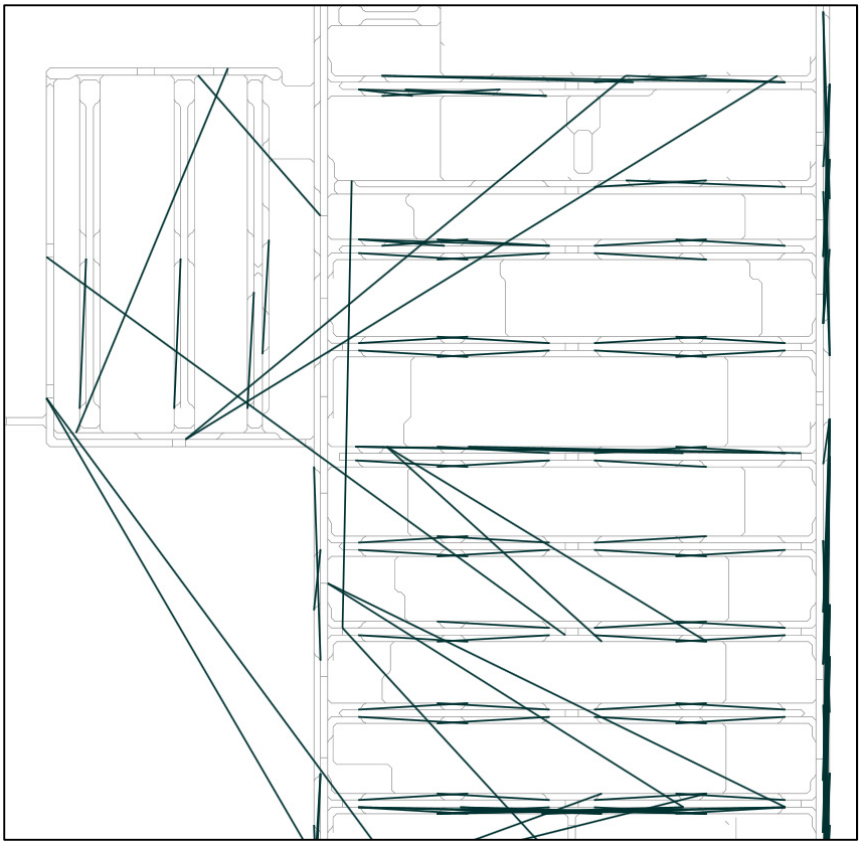

Figure 4: *Section of AMHS track layout of one of GLOBALFOUNDRIES' Dresden fabs (gray) with marked intersection pairs with multiple shortest paths (black)*

While in the real system cost parameters are limited to discrete values originally resembling length rounded to decimeters, in simulation models this limit does not apply. It is assumed that in reality as well as in simulation the path taken for one origin-destination relation is static, no splitting among equidistant paths takes place. To assess the potential impact of equidistances in this system cost parameters were adjusted in simulation runs slightly by small fractions of minimum differences of the real system (similar to Fig. 2). By this equidistances were eliminated, all shortest paths were made unique without changing the routing in general. Several of these cost parameter settings (each with other unique shortest paths) were generated, each showing one potential outcome of the actual real settings with multiple shortest paths; i.e. the ambiguous real cost parameters may lead to any of the unique

behaviors induced by these parameter settings adjusted for simulation, it is just not possible to predict which one is in fact the result in the real system. The simulation results (Tab. 1) therefore show a range of potential system behaviors and performance measures initiated by one and the same cost parameter setting. The sole difference is the selection of one of the shortest paths.

Table 1: *Result range of simulation runs with varying settings for multiple shortest paths in terms of delivery times (DT) normalised to values of best scenario*

		Medium load	High load
Best scenario	Average DT	1.00	1.00
	95th percentile DT	1.00	1.00
Worst scenario	Average DT	1.01	1.05
	95th percentile DT	1.01	1.19

The results for medium load scenarios are quite similar: In case no traffic congestions occur the delivery time is barely affected as the track length is obviously nearly the same. However, this does not hold true for a high load scenario. The differences in average delivery time of 5 % and in the 95th percentile of delivery time of 19 % show that there have been substantially more congestions in the worst case than in the best one. In fact, to achieve similar delivery times for the worst case would need to lower the load by 3 %.

To further test the potential impact of non-unique shortest paths the scenario with cost parameters reset to rounded decimetres of track length was simulated without subsequent optimisation. The throughput limit of this scenario was naturally considerably lower. However, executing several runs with only changing selection in case of multiple shortest paths showed an even bigger range for this throughput limit: The limit of the worst case was 16 % below the one of the best case. The optimisation process mentioned above obviously already reduced the number and impact of non-unique shortest paths as the difference for the maximum throughput was much smaller then.

In the real system the equidistances had to be eliminated differently as the limitation to discrete values resembling lengths rounded to decimetres applies. Therefore, the approach presented in section 3.3 had been employed in combination with the original optimisation algorithm. The two methods were executed in alternation. Thus, results of the respective simulation runs nearly identical to the ones of the best case in Tab. 1 could be achieved. In this scenario there are no ambiguities present in routing anymore. Considering and avoiding multiple shortest paths for origin-destination relations hence paid off in terms of predictability of system behaviour by simulation models as well as usability of simulation for system optimisation.

5 Conclusion

Even though shortest paths (regarding cost parameters) are often employed in some way as routing scheme for automated material handling systems, the case of non-

unique shortest paths is mostly ignored in both the real system and the simulation model. Explicit rules hence do rarely exist. Incorporating the same implicit rules in simulation models as taking effect in the real system is therefore hardly practicable.

For a real world case study it has been shown that the results of simulation studies in terms of delivery times and maximum throughput may vary in a range of up to more than 10 % relative to the best scenario. This makes simulation studies as prediction tool for the impact especially of different scenarios for cost settings affecting the routing unsuitable.

To assess this effect a method has been presented to identify a minimum set of origin-destination relations for which multiple shortest paths do exist. Out of this approaches have been described eliminating the occurrence of such equidistances with no or little impact to the general routing, depending on degree of freedom for cost parameters.

The importance of considering equidistances in a real system has been proven in a case study. Furthermore, the functionality of the described methods has been demonstrated providing a setting without any equidistances but achieving performance measures equal to the best case scenario without eliminating equidistances.

References

Bartlett, K.; Lee, J.; Ahmed, S,; Nemhauser, G.; Sokol, J.; Na B.: Congestion-aware dynamic routing in automated material handling systems. Computers & Industrial Engineering 70 (2014), pp. 176-182.

Dijkstra, E.W.: A note on two problems in connexion with graphs. Numerische Mathematik 1 (1959) 1, pp. 269-271.

Floyd, R.W.: Algorithm 97 Shortest Path. Communications of the ACM 5 (1962) 6, p. 345.

Gaxiola, G.; Christensen, E.; Hammel, C.; Stachura, P.: Methodology to best extend AMHS for site expansion. In: Laroque, C.; Himmelspach, J; Pasupathy, R.; Rose, O,; Uhrmacher, A.M. (eds.): Proceedings of the 2012 Winter Simulation Conference, Berlin (Germany), 9-12 December 2012, pp. 1967-1976.

Hammel, C.; Schmidt, T.; Schöps, M.: Network optimization prior to dynamic simulation of AMHS. In: Laroque, C.; Himmelspach, J; Pasupathy, R.; Rose, O,; Uhrmacher, A.M. (eds.): Proceedings of the 2012 Winter Simulation Conference, Berlin (Germany), 9-12 December 2012, pp. 1956-1966.

Warshall, S.: A Theorem on Boolean Matrices. Journal of the ACM 9 (1962) 1, pp. 11–12.