

YONG CUI, Ph.D.

E-mail: yong.cui@ievwwi.uni-stuttgart.de

ULLRICH MARTIN, Ph.D.

E-mail: ullrich.martin@ievwwi.uni-stuttgart.de

University of Stuttgart,

Institute of Railway and Transportation Engineering

Pfaffenwaldring 7, 70199, Stuttgart, Germany

Traffic Planning

Review

Accepted: Oct. 15, 2010

Approved: Nov. 15, 2011

MULTI-SCALE SIMULATION IN RAILWAY PLANNING AND OPERATION

ABSTRACT

Simulation methods are widely used in railway planning and operation. However, at the moment there are no applicable solutions in the process simulation for a smooth transition among different infrastructure levels on the basis of a unified structure with consistent algorithm. In this paper, a multi-scale simulation model is designed with consideration of the level of detail of the investigated infrastructure model and the homogeneity of the processes running in the simulation model. A comprehensive and synthesized view of railway planning and operation is therefore obtained. Within the multi-scale simulation model, railway planning and operation processes can be simulated, evaluated and optimized consistently.

KEY WORDS

railway planning, simulation, multi-scale, aggregation, discrete scaling, continuous scaling, homogenous process, inhomogeneous process

1. INTRODUCTION

In railway planning and operation, it is a common requirement to generate and calculate the information of operation processes in order to create timetable and/or evaluate the effects of a certain operation program. Some possible scenarios of using simulation tools can be the following:

- timetable construction for railway infrastructure companies and railway operators;
- generation of a new real-time dispatching timetable in case of deviation, e.g. in railway operations control center;
- capacity research;
- design and evaluation of logistic operation concepts and infrastructure layout.

With manual or analytic methods, it is inefficient or even impossible to determine the sophisticated train movements as well as the interaction among trains and infrastructure (signals, routes, and interlocking

system) for a complex network. Today, the simulation methods are widely used by railway planners, operators, and researchers.

By simulating the train movements and logging the information during the whole simulation process, the information of the given operation program can be obtained and evaluated. A round of simulation is similar to an experiment of a certain operation program. In order to get the statistic characteristics of certain operation programs, or to find a relative optimal timetable in a certain time period, several rounds of simulation may be executed with different input parameters and operational conditions. Sometimes the computational workload is huge. Therefore, the accuracy, the performance and the effectiveness of the utilized simulation model should be carefully assessed and balanced.

The design of a simulation model relies highly on the level of infrastructure detail, which can be categorized as *macroscopic*, *mesoscopic*, and *microscopic*. A macroscopic simulation is utilized for a large scale of simulation territory, which may cover the entire network, a subnet, a complete line, or a large area. Inside a macroscopic model, the infrastructure will be modelled at the highest abstraction level in order to obtain meaningful simulation results with reasonable computational effort and time based on a rough database mostly. While macroscopic simulation is not able to derive detailed information of the simulated operation processes from an aggregated model, the microscopic simulation is applied for the application with the highest level of detail. The results of a microscopic simulation may relate to an individual route node, tracks, junctions, stations, a complete line, or a very small network. Although a microscopic simulation is capable to simulate operation processes in detail, it can only be used in a limited area due to the considerable computational effort. Therefore, the mesoscopic models are introduced as complements. Although in some application contexts mesoscopic models integrate both macroscopic level and microscopic level, in

this paper the mesoscopic model is only designated for the level of detail in the middle of microscopic and macroscopic level. More introduction and discussion of the railway simulation and infrastructure modelling are given in [1] and [2].

With current solutions, a simulation model is bounded with a certain infrastructure level, wherein following deficiencies exist:

- There is no smooth transition from macroscopic models to microscopic models in a consistent way. Since a simulation tool is limited to be used at a certain level, additional efforts (for example, combination of two different models/tools) should be invested for the applications working on multi-scale levels.
- It is still the lack of a unified structure for the different level model, on which a unified algorithm, such as evaluation methods, simulation workflow, or route search algorithm, can be utilized.

In this paper, the general components and different modes of simulation models are introduced in Section 2. Then the design of new approach of multi-scale simulation model is discussed in Section 3. Finally, there is the introduction of the practical use with software solutions and further possible development.

2. OVERVIEW OF A SIMULATION MODEL

A simulation model can be described from two perspectives: the structure perspective and the behaviour perspective. The structure perspective concerns the components for railway simulation, and the behaviour perspective reflects the workflow and mechanism of the simulation model.

2.1 Components in a simulation model

In a simulation model, the railway infrastructure, rolling stocks, as well as operation program are respectively modelled as simulation components: infrastructure resources, performers, and simulation tasks.

In a multi-scale simulation, the infrastructure resources can be track segments, block sections, or aggregated resources including route nodes, track groups, stations, junctions or lines connecting stations and/or junctions. An infrastructure resource can be requested and allocated by a performer. The performers are the vehicles moving into, within, and/or out of the observed network, wherein each performer is assigned a simulation task. A simulation task describes the movements expected to be completed for a simulation performer in the operation. In a simulation task at least one route with a starting location and a destination location is defined. The definition of the route and starting/destination location is associated with infrastructure resources. The simulation tasks are

created based on an operation program, which also includes departure time and scheduled dwell time. For the trains scheduled with fixed time interval, several simulation tasks and performers can be automatically generated from a predefined operation program and train pattern.

After a simulation process is completed, the operation processes can be derived from the logged information. Further investigation and evaluation (for example, capacity research) can be carried out from the simulation results. It is also possible to visualize and monitor train movements during the simulation process when a synchronous simulation is applied (see Section 2.2).

2.2 Synchronous simulation or asynchronous simulation

Simulation methods can be classified as synchronous simulation and asynchronous simulation. In a synchronous simulation, all train movements are simulated simultaneously. Train schedules are constructed gradually driven by each time interval. Synchronous simulation is suitable to demonstrate and visualize train movements. The workflow of synchronous simulation (see Section 3.1) is very clear and easy to be implemented. The impacts of the deviation, which can be modelled as different types of delays in the model, are simulated similar to the reality as in railway dispatching systems. But in synchronous simulation, deadlocks may happen in the lines or in the station limits with bidirectional operations (see Section 3.2).

In contrast to synchronous simulation, deadlocks do not take place in an asynchronous simulation. Normally, trains are ranked according to their priorities in an asynchronous simulation model. Each train is introduced into simulation process successively based on the ranks. Since the complete blocking time stairway of a train was constructed before the train is inserted into the simulation model, all the train movements are ensured to be completed their path without deadlock problems.

With pre-built blocking time stairway, a train can be scheduled very fast in an asynchronous model. In addition, the time and efforts for deadlock avoidance are saved. The buffer time among the involved trains can be simulated close to the reality. The disadvantage of asynchronous simulation is the inflexibility of the fixed blocking time stairway, which may very often vary in real operation. The time and efforts to process possible conflicts and knock-on conflicts should be taken into consideration for some very busy lines. The pre-determined train priorities may violate the anti-discriminatory principle, if the priorities are defined based on a possibly biased foundation.

Synchronous simulation and asynchronous simulation can be combined and incorporated with each

other (see [1] and [3]). The precondition of asynchronous simulation, construction of blocking time stairway, can be fulfilled by synchronous simulation. Some trains with absolute higher priority can be simulated first, and the rest of the trains are introduced afterwards. The arrangement of different train groups fits the characteristics of asynchronous simulation, and the simulation process of each train group can be regarded as a synchronous simulation. Therefore, synchronous simulation may be considered as the basic form of simulation.

3. NEW APPROACH OF MULTI-SCALE SIMULATION MODEL

As the basic form of simulation, synchronous simulation can be utilized in a multi-scale simulation model. The workflow of synchronous simulation and associated deadlock problems will be discussed in this section. The most challenging issue of developing a multi-scale simulation model is to scale and transit in the simulated infrastructure model smoothly with different levels of detail. A new hybrid model has been designed, which integrates the processes among microscopic, mesoscopic, and macroscopic models consistently.

3.1 Workflow

The workflow of synchronous simulation reflects the behavioural features of the system. The diagram of the workflow for synchronous simulation is shown in Figure 1. The processes of a synchronous simulation are triggered by certain fixed or varied time intervals. Inside each time interval, three activities – request resources, allocate resources, and proceed with simulation tasks – are executed.

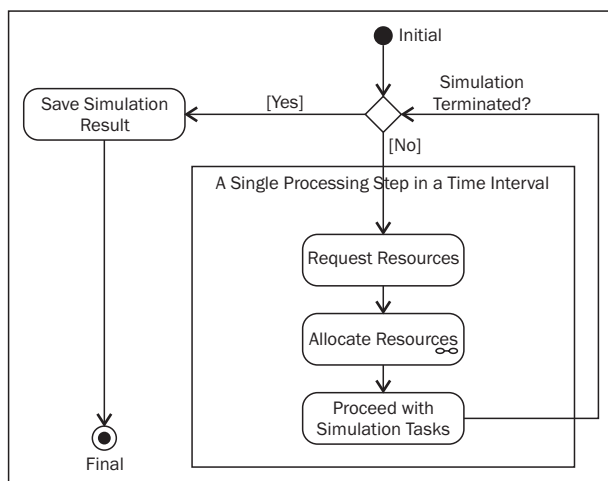


Figure 1 - The workflow of synchronous simulation

Source: [6]

In each time interval, each train forecasts and requests the infrastructure resources supposed to be entered in the said interval. The requests are validated through conflict-free tests and deadlock-free tests. If an infrastructure resource can be entered by multiple requesters (the request associated to each requester is conflict-free and deadlock-free), certain assessment rules are utilized to determine which performer can get the chance to occupy the said resource. After the allocation of the requested resources is completed, the position of trains and the status of infrastructure are updated to simulate the change of the system in the current time interval.

The workflow of synchronous simulation is clear and intuitive. Many simulation tools are developed based on the workflow. However, a synchronous simulation model has to deal with the challenges from deadlock problems.

3.2 Deadlock problem

A deadlock is defined as follows [4]: “... a situation in which a number of trains cannot continue their path at all because every train is blocked by another one.”

Although many deadlock situations may be manually identified and avoided in real operation, it can frequently happen in the computer synchronous simulation on the tracks with bidirectional operation. In the example shown in Figure 2, without recognizing other concurrent train movements, Train A enters into Track 3 and therefore leads to deadlock.



Figure 2 - An example of deadlock

The research for resolving deadlock problems started at the end of 1960s, originally applied for computer sciences. Three strategies – deadlock detection, deadlock prevention, and deadlock avoidance – are developed. Only deadlock avoidance is suitable for railway operation and simulation [4].

Several deadlock avoidance solutions are developed and modified for railway operation and simulation, including movement consequence analysis (MCA), dynamic route reservation (DRR) [4], the banker’s algorithm [5], the Peterson and Taylor algorithm [7], and the labelling algorithm [8]. However, a universal and practical solution to completely solve the deadlock problems is not available yet. An essential problem for the deadlock avoidance algorithm is associated to determining how far in advance of the operation processes should be considered. On the one hand, it may still lead to deadlocks due to a too short investigation space. On the other hand, many situa-

tions that not necessarily lead to deadlocks may be identified as deadlock positively and the usability of the solution is therefore limited, if the investigation space is too large. The core of many solutions is to define practical rules [4] and improvements [5] in order to find a suitable boundary.

3.3 Scaling and link of scales

Scaling of infrastructure model includes two directions: from microscopic level to macroscopic level (bottom-up) or the opposite (top-down) [2].

In the bottom-up direction, the attributes are aggregated from the microscopic level to upper levels. Some aggregated attributes, for example, the number of trains that passed in a node, or the overall waiting time for a small area can be easily derived from the microscopic level. However, not all the attributes can be determined in a clear and unambiguous way. For instance, when calculating the maximal permitted speed of a certain railway line, it can be the lowest permitted maximal speed of a segment in the examined line, or the average permitted speed weighted by the length of the segments. Therefore, it is necessary to define a context-dependent algorithm for each application to aggregate the indicators from the microscopic level to upper levels.

It is hard to derive the value of attributes from an abstracted macroscopic level to a detailed microscopic level. In a top-down direction, additional detailed information need to be estimated and extrapolated from the aggregated data. Therefore, the usability of the derived data is limited by the quality and accuracy of the model. Furthermore, the effort to build a model and to calibrate the parameters of the model is also considerable. A top-down scaling is not widely utilized in practise yet.

For a multi-scale simulation model, it is meaningful to apply a unified graphical structure for each level, so that a general simulation workflow, algorithm, or evaluation method can be applied consistently. In railway infrastructure modelling, a very practical model is the link and node model [2]. The possible link and node model for different levels is shown in Table 1.

Table 1 - The link and node model for different infrastructure levels

Infrastructure Level	Node	Link
Macroscopic	Station, Junction	Line between Station/Junction
Mesoscopic	Route Node	Track, Track Group,
Microscopic	Points, Crossing, Signal	Track, Block Section

The link of the different level can be achieved through the aggregation model. Each node or link at

an aggregated level can be regarded as the set of the nodes or links in its lower level. For example, a station consists of several route nodes and track groups, wherein each route node consists of several points, signals and tracks, and each track group consists of several tracks, signals, and block sections. The hierarchy of the links is shown in Figure 3.

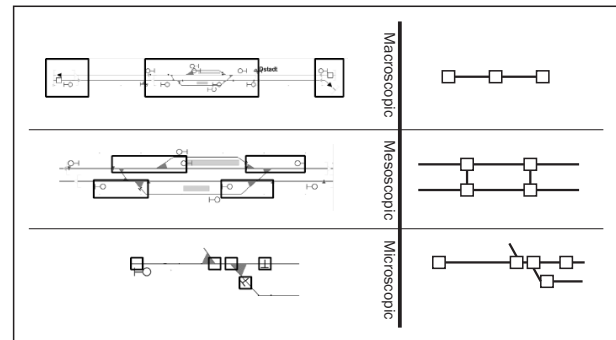


Figure 3 - The aggregation of infrastructure link and node model

The scaling of the simulation model can be *discrete* or *continuous*. In a discrete scaling mode, the user can either explicitly specify a level, or shift to a certain level according to a pre-defined threshold (for example, the zoom-in/out rate). The discrete model is easy to be understood and implemented. However, it can't provide a smooth and transparent transition between different levels. Sometimes, the explicit scaling is implemented through switching between different software tools, which is also inconvenient for the users.

Continuous scaling provides a synthesized platform for infrastructure presentation and multi-scale simulation. An assessment model is required to evaluate the significance of a node/link. The level of detail of a node/link processed in the model will be determined by the assessed significance value. The criteria for the assessment depend on the type of investigation and expected results. Some examples can be:

- the total length of the evaluated node/link
- the number of points/crossing/block section
- the number of route node/track groups
- the type of services (passenger, freight, mixed, or shunting)
- the capacity of the evaluated node/link
- other user-specified values.

Every infrastructure node/link can get at least a chance to be processed at an aggregated macroscopic level. The value of the significance and the current zoom-in/out rate determines whether the evaluated node/link can be further processed in a mesoscopic or microscopic level. With a low zoom-in rate, the nodes/links with higher significance may be presented and simulated at a more detailed infrastructure level than those with lower significance. With the increase in zoom-in rate, more and more details can be processed in detail. Therefore, a continuous scaling of simulation

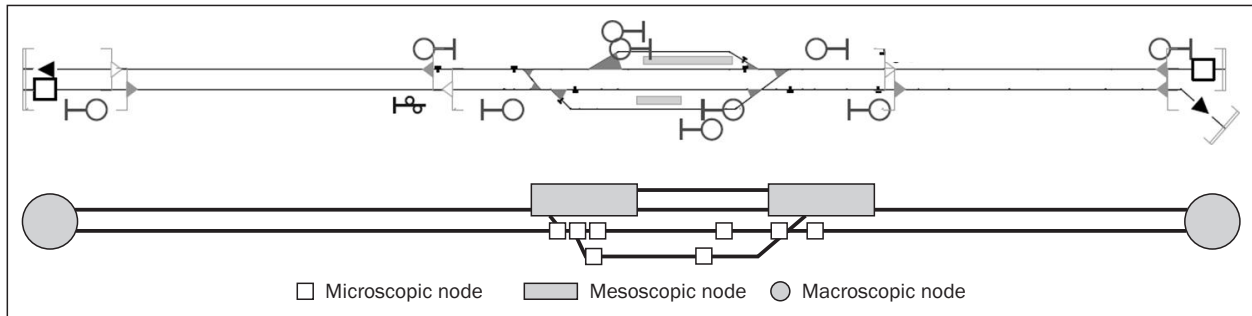


Figure 4 - An example of continuous scaling mode

model is achieved. In *Figure 4*, an example of continuous scaling mode is illustrated at a certain zoom-in/out rate. According to the value of the significance, all the infrastructure nodes are displayed and processed at microscopic, mesoscopic, or macroscopic level, respectively.

3.4 Design of a new hybrid model

A simulation model is designed for one or more railway planning or operational purposes. Based on the purposes, some specific processes will be executed. These processes can be:

- scheduling of operation processes;
- dispatching;
- optimization;
- capacity research;
- visualization of train movements and operation processes.

A hybrid model has therefore been designed, which integrates the processes from microscopic level to macroscopic level together with inside a common process-oriented framework. According to the characteristics of the processes, either the discrete scaling mode or continuous scaling mode (see Section 3.3) can be chosen.

The discrete scaling mode is suitable for the scenario, where the purpose and the workflow of the simulation processes for different infrastructure levels are distinct. Such processes can be regarded as *inhomogeneous* processes. An example of the framework is the hybrid model of multi-level dispatching framework proposed in [6]. Inside the framework, a detailed synchronous simulation is executed at the microscopic level to derive the basic dispatching solution. The basic solution will be aggregated and optimized at the macroscopic level. Finally, the optimized solution will be further elaborated at the microscopic level in order to obtain the final dispatching solution. In the said framework, the processes at microscopic level (scheduling, elaboration) and the process at macroscopic level (optimization) have different workflow and intention. By isolating the microscopic details from the strategy optimization process different concerns are encapsu-

lated at different levels. The system flexibility is gained with low level of complexity.

For the research that not only requires the perspective from the macroscopic level, but also concentrates on the details in one or more local areas simultaneously, a hybrid model can be designed based on the continuous scaling mode. Inside this model, one or more processes can run at both macroscopic and microscopic level. Such processes can be regarded as *homogenous* processes. For example, through the simulation running on the continuous scaling mode, the quality of service (such as waiting time or the grade of punctuality) and the utilization level of infrastructure nodes/links can be calculated and evaluated at the microscopic, mesoscopic, and macroscopic levels concurrently. The observed areas, bottlenecks, or the areas with high significance, will be simulated and examined at the microscopic level. The rest areas are investigated at mesoscopic or macroscopic levels with a low level of detail.

The homogeneity of processes should be determined at the time to design and construct the simulation model. If a process is embodied inside the simulation model, it is favourable to design the process as a homogenous process (e.g. a route searching process), which can be shared and applied for different infrastructure detail levels with a consistent workflow. However, not all processes are suitable to be designed or regarded as homogenous processes. Situations using inhomogeneous processes are:

- The computational efforts of a process are huge with an unnecessarily high level of detail. The process is suitable to be considered in a high aggregated level only.
- Due to the high requirement of level of detail, it only makes sense to design some processes in a microscopic level (e.g. the optimized energy-efficient driving style and running dynamic calculation).
- If a simulation model is designed for multiple processes (e.g. evaluation, operation control, or analysis), which are running outside the application context of the simulation model or in a third-party software with different specific level of detail, these processes can be considered as inhomogeneous processes.

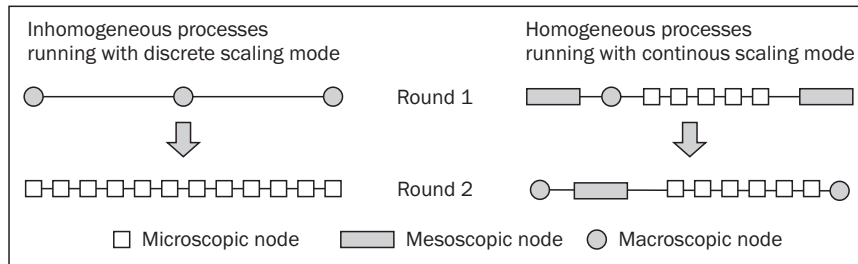


Figure 5 - The comparison of the multiphase simulation with different scaling modes

In a hybrid simulation model, processes can be executed with several rounds. For inhomogeneous processes, a round of simulation at a certain infrastructure level is designated for a certain phase that is suitable for the current level. For homogeneous processes, a round of simulation covers both macroscopic and microscopic levels. Several rounds of simulation for homogeneous processes can be used for a multiphase workflow and/or to derive an optimized result. In each round, the concerned level of detail of the infrastructure nodes/links may vary dynamically according to the new assessed significance value. The comparison of the multiphase simulation for different scaling modes is shown in *Figure 5*.

Scaling of the models concerns on the level of detail of infrastructure models, and hybrid model concerns the processes executed in the model. The homogeneity of the processes determines which kind of scaling mode is applied. A discrete scaling mode is suitable for inhomogeneous processes; homogeneous processes are preferred to be simulated in a continuous scaling without needing switches between different infrastructure levels.

4. PRACTICAL USE WITH SOFTWARE SOLUTIONS AND FURTHER DEVELOPMENT

Some railway simulation software tools are available for commercial and research purposes. They are based on either synchronous mode, such as RailSys (by RMCon Software), and OpenTrack (by ETH Zurich), or asynchronous mode, such as LUKS (by RWTH Aachen). These tools can be used for long-term or short-term timetable construction, capacity research, and operation control. It can also be used for design and evaluation of logistic operation concepts, such as the software tool PULRAN (by University of Stuttgart), which has been developed to simulate the shunting process with the feature of deadlock avoidance.

The current solutions for multi-scale simulation are mainly implemented by combining macroscopic and microscopic simulation through a discrete scaling model. An example is the combination between RailSys (microscopic simulation) and NEMO (macroscopic, by RMCon Software). To achieve a smooth and con-

sistent transition with a hybrid simulation model, it is efficient to unify the processes running on different infrastructure levels as much as possible. Further development of multi-scale simulation can be the design of a continuous scaling mode applied with consistent processes in the model, wherein the assessment of the significance of infrastructure node/link is decisive.

Since it is hard to derive the detailed information of infrastructure from macroscopic to microscopic, data aggregation is more practical for transformation through different infrastructure levels in a multi-scale simulation model. Further development of a multi-scale simulation should start from continuous data collection including an iterative consideration at microscopic level, and the design of infrastructure aggregation model.

To verify the usability of a simulation model, the results of a simulation model can be compared with the reality. The quality and the effectiveness of simulation tools used in railway planning and operation will be further improved parallel with the development of the railway system modelling, infrastructure data management, as well as the research and the application of multi-scale simulation models.

5. CONCLUSION

In railway planning and operation, simulation methods are used as experimental approaches for timetable construction, evaluation of infrastructure designs and operation concepts, as well as train dispatching. Two different modes of simulation, synchronous and asynchronous simulation, can be incorporated with each other.

With current solutions, a simulation model is bounded with a certain infrastructure level. The efficiency and effectiveness are limited due to lack of a unified structure and algorithm among different infrastructure levels, as well as a smooth transition between levels. By differentiation of the level of detail of the investigated infrastructure, a flexible transition using discrete or continuous scaling will be achieved with consideration of the homogeneity of the processes running in the simulation model. Within the multi-scale simulation model, railway planning and operation processes can be simulated, evaluated and optimized consistently.

Dr.-Ing. **YONG CUI**

E-mail: yong.cui@ievwwi.uni-stuttgart.de

Dr.-Ing. **ULLRICH MARTIN**

E-mail: ullrich.martin@ievwwi.uni-stuttgart.de

Universität Stuttgart,
Institut für Eisenbahn- und Verkehrswesen
Pfaffenwaldring 7, 70199, Stuttgart, Germany

ZUSAMMENFASSUNG

MEHRSKALEN-SIMULATION IN EISENBAHNBETRIEB UND INFRASTRUKTURPLANUNG

Simulationsmethoden sind bei der Gestaltung des Eisenbahnbetriebs und der Infrastrukturplanung weit verbreitet. Dabei werden unterschiedliche Detaillierungsgrade der Betrachtung genutzt. Derzeit gibt es jedoch noch keine anwendungsorientierten Lösungen bei der Prozesssimulation für einen reibungslosen Übergang zwischen den verschiedenen Infrastrukturebenen auf der Grundlage einer einheitlichen Struktur mit konsistentem Algorithmus. Nachfolgend wird ein Mehrskalen-Simulationsmodell vorgestellt, das die unterschiedlichen Detaillierungsgrade der Infrastruktur berücksichtigt und die Homogenität der Prozesse bei der Simulation sicherstellt. Der vorgestellte Ansatz ermöglicht eine umfassende ganzheitliche Betrachtungsweise des Planungs- und Betriebsprozesses bei der Simulation, Evaluierung und Optimierung.

SCHLAGWORTE

Eisenbahnplanung, Simulation, Mehrskalen, Aggregation, diskrete Skalierung, durchgängige Skalierung, homogene Prozess, inhomogene Prozess

LITERATURE

- [1] **Martin, U.** (co-author): *Fahrwegkapazität: Leistungsuntersuchungen mit Simulationsmethoden*, DB Netz AG Rechtlinie 405.0105, Frankfurt, 2006
- [2] **Radtke, A.** (co-author): *Infrastructure Modeling*, In *Railway Timetable & Traffic*, Eurailpress, Hamburg, 2008
- [3] **Jacobs, J.** (co-author): *Conflict Resolution*, In *Railway Timetable & Traffic*, Eurailpress, Hamburg, 2008
- [4] **Pachl, J.**: *Avoiding Deadlocks in Synchronous Railway Simulations*, 2nd International Seminar on Railway Operations Modeling and Analysis, Hannover, March 2007
- [5] **Martin, U., Cui, Y.**: *Method and System for Simulating, Planning, and/or Controlling Operating Processes in a Track Guided Transportation System*, European Patent Application No./Patent No.: 10006446.8 - 2215, Munich, August 2010
- [6] **Cui, Y.**: *Simulation-Based Hybrid-Model for a Partially-Automatic Dispatching of Railway Operation*, Ph.D. thesis, University of Stuttgart, Stuttgart, Germany, 2010
- [7] **Petersen, E. R., Taylor, A. J.**: *Line Block Prevention in Rail Line Dispatch*, *INFOR Journal* 21, No. 1, 1983, pp. 46-51
- [8] **Mills, G., Pudney, P.**: *The Effects of Deadlock Avoidance on Rail Network Capacity and Performance*, Mathematics-in-Industry Study Group Workshop, Adelaide, Australia, February 2003