Measuring the Performance Availability of a Logistics Application Controlled by a Wireless Sensor Network

Messung der Leistungsverfügbarkeit in einem über drahtlose Netzwerke gesteuerten logistischen System

Moritz Roidl, Steffen Schieweck, Michael ten Hompel, TU Dortmund, Dortmund (Germany), moritz.roidl@tu-dortmund.de, steffen.schieweck@tu-dortmund.de, michael.ten.hompel@iml.fraunhofer.de

Abstract: Performance availability as a Key Performance Indicator (KPI) for facility logistics systems has become a topic of interest in recent years. It deals with the computation of an entire system’s performance which is influenced by the availability of its subsystems. The importance of high-quality planning of facility logistics systems also increased dramatically regarding cost reduction and heightened demands. Simultaneously, decentralised logistics systems controlled by wireless sensor networks have taken a large step towards industrial application. An approach to measure performance availability in systems influenced by the effects of wireless sensor networks is presented here. A brief case study is provided to illustrate the effects of wireless sensor networks on performance availability.

1 Introduction

The two core elements that provide understanding of the desired performance of facility logistics systems are maximum performance and share of time at which the measured performance can be provided. Traditionally these two indicators have been tested separately at the launch of a facility logistics system (Maier 2012). In recent years, the concept of performance availability has been established which aims to combine both indicators. For this purpose, the VDI 4486 (2012) guideline provides handover testing procedures after the launch of the system as well as required definitions and agreements. The guideline is meant to be part of the contractual agreements between the vendor and the operator of facility logistics systems.

Typically, for the forecasting of facility logistics systems KPIs, a system gets subdivided into several manageable and isolated subsystems for which proven methods are available (ten Hompel et al. 2007). Consequently, interdependencies which occur between those subsystems are neglected. However, this may lead to insufficient results for strongly connected material flow systems, especially when
considering a system-focused KPI such as performance availability. Event-discrete simulation is highly suitable for the investigation of the timely behaviour of complex material flow systems (März et al. 2011). Though, the creation of simulation models is a complex, labour-intense and thus an expensive task.

For distributed systems, with resource constrained devices, the communication behaviour significantly influences the overall system performance. From the perspective of a classical logistics planner, the evaluation of a system would be merely a question of technical availability of the individual components. Nevertheless, there is currently no approach towards specific performance requirements for planning logistics systems that include wireless sensor networks.

A tool is presented here capable of measuring performance availability of material flow systems. To begin with, the basics of performance availability are described in chapter 2. Chapter 3 presents the implemented procedures for performance availability measurement. Special requirements for systems with wireless sensor networks are explained in chapter 4 subsequently the architecture of the tool is described in chapter 5. The results of some experiments comparing conventional systems to systems with wireless sensor networks are depicted in section 6. The article concludes with a brief summary and an outlook.

2 Performance Availability

Performance availability as a KPI has been established in the recent years. Practitioners agree that its principles strongly meet existing demands. In this section, an overview of the ongoing research is presented. A survey of relevant literature describes the definition of performance availability which is applied in this article.

2.1 Literature Overview

Before the term performance availability was established, research mainly focused on the technical availability of an entity, “the probability of finding the observed units in a correctly functioning state at any given time during the operating period” (VDI 3581 2004, p. 2). The performance of a facility logistics system, in the context of this paper, is defined as the throughput of material flow objects of either the whole facility logistics system or its subsystems or entities. Performance availability has first been brought to context by Wittenstein (2007) who establishes the term in the domain of customer specific product development in systems and plant engineering as a state in which a process is conducted according to its specifications. Hence, required resources need to be available at any time independent of disturbances or fluctuations in demand. The focus of the article lies on the design of a structured procedure for the fulfilment of performance as defined by a customer in development assignments. Maier (2012) proposes practice-oriented acceptance procedures regarding reliability and availability standards, considering the analysis of existing methods and thus deducing recommendations. She concludes that coexisting standards are not or only partially suitable for facility logistics systems. The work lays the foundations for the VDI 4486 (2012) whose definition of performance availability will be used for the remainders of this article and is described with further detail in section 2.2. It provides standard procedures for the measurement of performance availability of existing facility logistics systems.
The VDI 4486 (2012) guideline has served as a landmark for succeeding publications, being the first official standard to deal with performance availability. Based on a preliminary version of the VDI 4486 (2012), Maier (2010) proposes an analytical method for the forecasting of performance availability. A more detailed approach to that matter is given by Jung and ten Hompel (2013) who aim to determine the number of delayed objects when passing a branch in a conveyor system. Hegmanns et al. (2014) presents a concept for the determination of performance availability with varying focuses. In this concept, a number of analytical methods and simulation models are used to cover different aspects of performance availability for different material flow technologies. Some principles of simulation modelling for the measurement of performance availability are proposed by Schieweck et al. (2014). Aiming to push the research towards practical applicability, Schieweck and Jung (2015) discuss the demands of practitioners and the current degree of fulfilment provided by the published research.

2.2 Definition

As stated in the previous section, the VDI 4486 (2012) guideline is considered the prevailing standard definition of performance availability. Hence, its definitions and measurement procedures are described briefly in this section. “The performance availability indicates the degree of fulfilment of processes agreed between contract parties (manufacturer and user) in accordance with the requirements and deadlines and in compliance with the agreed basic conditions” (VDI 4486 2012, p. 3). Thus, performance availability is defined as the degree of fulfilment of specified business goals. Moreover, measurement processes are defined which add a quantitative dimension to the term which is required for a verifiable contractual agreement.

The measurement as defined by VDI 4486 (2012) requires the definition of business goals. Those need to be either measurable in waiting times or the timeliness of the fulfilment of orders (VDI 4486 uses the term “running times”). Examples for business goals are the utilization of pickers or the service level of a picking process. When considering waiting times the degree of fulfilment is given by the ratio of the overall time $T_B$ minus the waiting time $T_W$ and the overall time $T_B$ (VDI 4486 2012, p.6).

$$\eta_w = \frac{T_B - T_W}{T_B} \quad (1)$$

It must be noted that the measurement of performance availability needs to be conducted in several observation periods for each of which the fulfilment of performance specifications is checked. In any case the desired performance was provided within the observation period, the performance availability value is set to 100%. For any other case the aforementioned formula is used. For the determination of the overall performance availability the average of the considered observation periods is calculated. The consideration of timeliness requires the counting of the overall number of orders $N$ and the delayed orders $n$ at the time of completion (VDI 4486 2012, p.7).
\[ \eta_t = \frac{N - n}{N} \]  

(2)

As before, first the achievement of specified performance needs to be checked for the observation period before applying the formula. Also, the overall performance availability is calculated as the average of the relevant values for the observation periods. The real-world acceptance procedure not only requires the specification of performance values but also the time in which a disturbance within the system needs to be resolved. Further conditions and agreements for the acceptance procedures can be found in VDI 4486 (2012). The guideline can also be referred to regarding useful examples and more formal definitions.

3 Performance Availability Measurement

Basic principles of the measurement of performance availability have been discussed by Schieweck et al. (2014). In this section, those principles will be concretized by describing the procedures implemented to the designed simulation tool. To begin with, the necessary input and output data are defined. Following, a description of the implemented structure for the proposed tool is provided.

3.1 Data

The defined data are subdivided into the categories of input and output data. The input data are necessary for the simulation and need to be created before the simulation runs. The output data are gained during the simulation runs, exported, analyzed and interpreted. Necessary input data sets in terms of the evaluation of performance availability are:

- picking list
- initial inventory
- inflow
- due dates

All those data sets are part of a real-world facility logistics system. The picking list defines the load of the picking system and comprises a number of picking lines. Usually, one picking line consists of the time of order creation, order number, number of the picking line, article number, and pieces per article. The initial inventory describes the amount and type of articles which are in stock at the time of simulation start. The inflow describes the flow of goods which arrive at the warehouse and are to be stored in the system. The due dates are defined for every order and specify their latest time of completion. Their creation and definition have considerable effect on the performance availability value when considering timeliness. Commonly available for simulation studies in industrial projects are the former three. An approach for the creation of practically applicable due dates for facility logistics systems is provided by Schieweck et al. (2014). Necessary output data for the assessment of performance availability are:

- waiting time
- throughput
- creation time
• completion time

The collection of the aggregated waiting time of a resource is necessary for the first of the aforementioned definition of performance availability (see formula 1). The throughput for every observation period is required as well, following the definition in section 2.2. For the assessment of timeliness, creation and completion time of every order has to be recorded.

3.2 Tool Structure

The developed simulation tool aims to use the advantages of two available programs, namely Microsoft Excel and Demo3D by Emulate3D Ltd. The structure of the warehousing system can easily and quickly be defined in Microsoft Excel by inserting characters and tokens into the cells. The entire warehousing system from receiving to shipping area may be modelled. The structure is then imported to the environment of Demo3D in which the model is instantly able to operate and create results. The results are automatically fetched and exported, again to the common environment of Microsoft Excel. Using a predefined folder structure, a user is able to define several scenarios with variable layouts and/or data which can be tested one after the other without any further interaction. A considerable reduction of model-building and testing time for the measurement of performance availability can be achieved.

The creation and/or preparation of the necessary input data depends highly on the individual situation and is therefore not treated within this article. The existing data needs to be brought to a standardized format. After the data was inserted to the tool, it automatically creates CSV-files which keep all of the required data for further usage. Analogically, the defined layout and its specifications are exported from Microsoft Excel into CSV-files. Within Demo3D scripting objects import the CSV-files, create the defined layout within the 3D-environment and store the data in Demo3D-internal tables. Now, the model allows instantly for running and testing. A more detailed description of the simulation model’s internal functionality is given in section 4.

The recording of the data is coded within the individually created modules (e.g. conveyors, pickers) as well as in global scripting objects which handle the exporting of the data. The waiting time is recorded by each module individually. Every time a load on the module is willing to proceed but has to remain on the module the timeframe is recorded and added to the overall waiting time of the module. At the end of the simulation run a scripting object called “data fetcher” initiates a broadcast to all of the relevant modules requesting a waiting time report. Each module sends its own report. The fetched data is stored into a Demo3D-internal table and exported as an XLS-file. The measurement of timeliness can either be conducted internally in the model or externally in Microsoft Excel. When being conducted internally the method proposed by Schieweck et al. (2014) is implemented to the module which initiates the creation of order totes. Instantly, every order tote gets a due date assigned. At the shipping department, the due date is compared with the actual time and the order tote receives the tag “on time” or “delayed”. The data fetcher also requests timeliness reports from the shipping department. It must be noted that the waiting times and timeliness require the assessment of the achievement of the
throughput specifications before finally being able to be used for the calculation of performance availability.

The various scenarios are defined within scenario folders. The scenarios are identified with a distinct number. Each scenario has its own data files (input and output) as well as one layout file each. Within the tool, the user defines the run duration of each scenario and selects the scenarios which are to be tested by entering the scenario number. At the beginning of each scenario run, its individual layout is imported. During the scenario run the output data are collected and after the scenario has finished they are exported into Microsoft Excel tables within the associated scenario folder.

4 Wireless Sensor Networks in Facility Logistics Systems

The application of cyber-physical and embedded systems has been developed in the field of logistics during recent years as a part of the Internet of Things (IoT) revolution. Current research in the field of logistics investigates the massive use of distributed embedded systems as a new basis for modern production, transportation and distribution strategies. These efforts can be summarized under the heading Industry 4.0. It describes a concept that builds on a cyber-physical system comprised of individual smart devices taking over control from the production planning system. These devices organize the transport in a self-sufficient manner while using services of external systems along the supply chain. While the research in these decentralized controlled production and logistics systems has produced promising results in the recent years, the idea of a completely self-sufficient individual smart device has not yet found a convincing and practical implementation. Recent developments in low power technology for wireless sensor networks make it apparent that a breakthrough might be imminent.

For systems of distributed and resource-constrained devices the communication behaviour significantly influences the possible overall system performance. In accordance with the requirements for developing smart industrial devices adjacent to the IoT compliance, the inBin (intelligent bin) has been developed. Not only does it stores data about its contents and knows its position, it works in an energy-neutral manner. In addition to interaction possibilities with the operator through its display and buttons it includes a wireless transceiver that enables communication with its materials handling central management system.

In an industrial real-world scenario, a large number of inBins will work together simultaneously in the manner of a wireless sensor network within one facility. inBins should be able to communicate within the system to assure its functionality. The communication infrastructure is based on a central server which handles all devices and there is no direct inter-device communication. Except for a centralized management of transport orders, the comprehensive system functions are provided by inBins. This includes regular status updates from the nodes to the management system, query capability for specific articles, as well as distributed decision strategies for delegating a transport request to individual bins.

In contrast to other applications for sensor networks this architecture is feasible due to the highly controlled environments in facility logistics.
5 Simulation of Wireless Sensor Networks

The simulation proposed in this article evaluates the performance availability of a large number of inBins within a material flow simulation tool. As such, the main focus lies on the correct modelling of the application layer protocol which controls the material flow system in the manner of a multi-agent-based system. There are two main agent types in the system: the mobile inBins and the stationary material flow equipment. The inBins share a limited radio resource channel with a simple infrastructure by using a rudimentary communication protocol while the rest of the system entities (i.e. conveyors) are considered to be connected by cable. A single transceiver on the server side can reach all inBins. The communication protocol is based on time slots of uniform length that cover a complete request-acknowledge cycle with status data exchanged between the devices and the server.

The main characteristic of wireless communication in relation to performance availability is the unreliable timing of messages. The limited radio resource channel means that any time a specific inBin tries to send a message it can be delayed by messaging activity of other inBins. In addition, the mobility of the inBins restricts the energy that is available for active radio communication, so there can also be communication delays due to energy shortage. These delays may increase further when communication protocols are considered that represent a many-to-many relationship between inBins during multi-agent-based decision-making. In this case, one inBin might have to wait for the answer from an unknown number of inBins to be able to decide, for example, the most time-efficient route through a material-flow system. This can be handled by using timeouts that restrict the waiting time after which the decision will be made. The timeout has to be chosen sufficiently long to get enough answers for a decision and, thus, has to be based on the maximum expected latency. The more inBins try to answer the request the higher latency will be since the communication medium is shared.

The Demo3D-based tool was extended to include a network object through which all messages between agents (i.e. inBins and conveyors, etc.) are transmitted. The messages are modelled as property maps which include the sender and receiver and message type information, as well as the payload data. The network has a lookup table with all registered agents and offers two services as script functions: the sending of a message (one-to-one) and a message broadcast (one-to-many). These functions can be called by an agent in a non-blocking way, so that the message will be processed independently by the network while the agent program continues to run. The network then applies a latency to the message and holds the message property table until the delivery time has been reached. It then looks up the recipient of the message and delivers the message property table through a script function call on the receiving agent.

Based on this architecture, all logistics processes are modelled decentralized and communication-based in the simulation tool. One advantage of using the central network object is the possibility to visualize the messages within the simulation tool, and thus combining material and information flow into a unified view of the facility logistics system (see figure 1). This can help while debugging the logistics processes which depend on the correct interaction between the physical objects and virtual agents.
6 Case Study

The presented case study is aimed to demonstrate the functionality of the concept and presents some results about the impact of the network latency occurring in large logistics systems controlled by wireless sensor networks. First, the case study system and the data are defined. Following, some results will be presented and discussed.

6.1 System and Data

The case study system consists of a miniload storage area, a picking station, the connecting conveyor loop and a conveyor system for the order totes. The miniload system comprises 3 aisles with a storage capacity of 100 storage locations each. The overall storage capacity is 300 storage locations out of which 270 are occupied at the start of the simulation runs. The resulting storage utilization is 90%. At the picking station, each order line requires a time of 5 seconds for picking. The picking process may only be conducted when the required storage and order totes are at the desired positions at the picking station. The orders are defined as ‘completed’ as soon as the associated order tote arrives at the ‘shipping department’ which is shortly after the picking station.

Among the defined testing scenarios, the layout is not changed. For every hour of simulation time a mean of 50 orders are put into the system. The different scenarios are characterized by various sizes of the orders from 4 to 8 picking lines. All of the scenarios are tested incorporating a latency of 50 to 110 milliseconds for every message sent via the wireless sensor network. To assess the influence of the latency another set of testing runs are conducted with no latency at all. The due dates which are required for the investigation of the timeliness are created by the method proposed by Schieweck et al. (2014) with 4 routes, a departure pulse of 5 minutes and a critical value h which has a basic allowance of 40 and adds 5 seconds for every order line number greater than 4.
6.2 Results

For the assessment of the latencies influence two performance availability values are measured. First, every time a conveyor module requests a transport to the next conveyor module but has to wait due to missing capacity, its state is characterised as ‘blocked’. Every occurring timeframe is recorded as a waiting time as defined in formula (1). For the presented case the performance availability of the conveyor system is tested. To achieve this, the average value of the waiting time of the conveying modules is used then the timeliness of an order is checked. For the investigation of the performance availabilities behaviour, it is assumed that the specified performance is not reached at any point in time. Hence, all of the waiting times and tardy jobs are included to the calculation. The results are depicted in figure 2.

![Performance availability results](image)

**Figure 2: Performance availability results**

Generally, the performance availability of systems with considerable latency is lower than in systems with insignificant latency. This is easily explainable as the latency causes slower processes which result in larger waiting and throughput times. Looking at the performance availability of the conveyor loop this thesis can be supported. Even more crucial is the observation that with increasing load to the system (aka higher number of totes per order) the performance availability for scenarios with latency decreases disproportionately. This has special significance in facility logistics systems when the maximum performance is required while maximum system load appears. The performance availability based on the timeliness of the orders also supports the thesis of lower values for systems with latency.

7 Summary and Outlook

In this article an approach to measure performance availability in systems influenced by the effects of wireless sensor networks is presented. A brief case study based on a simulation tool developed in Demo3D provides an illustration of the effects and a proof-of-concept for the developed tool. Future research has to be conducted to investigate influencing factors on performance availability for several system types.
as well as the allowance on network latency for real-world systems. The latter has two key points of interest. Either the latency of a realistically sized decentralized system is measured or the maximum allowance for latency can be investigated for the achievement of satisfactory performance availability values. The simulation of real-world network latency might be conducted by simulation coupling (i.e. HLA).

**Acknowledgement**

The presented research is funded by the German Research Foundation DFG within the subprojects C1 and C5 which are part of the project PAK 672 (HO 2403/8-1).

**References**


