Monostori, L.; Ilie-Zudor, E.; Kádár, B.; Karnok, D.; Pfeiffer, A.; Kemény, Zs.; Szathmári, M.: Novel IT solutions for increasing transparency in production and in supply chains, Proceedings of the 21st International Conference on Production Research, ICPR 21, July 31 – August 4, 2011, Stuttgart, Germany, (ISBN: 978-3-8396-0293-5) paper no.: IS1.3, p. 6.

NOVEL IT SOLUTIONS FOR INCREASING TRANSPARENCY IN PRODUCTION AND IN SUPPLY CHAINS

L. Monostori[1,2], E. Ilie-Zudor[1], B. Kádár[1], D. Karnok[1,2], A. Pfeiffer[1], Zs. Kemény[1], M. Szathmári[1]
[1]Fraunhofer Project Center on Production Management and Informatics,
Computer and Automation Research Institute, Budapest, Hungary
[2]Department of Manufacturing Science and Technology,
Budapest University of Technology and Economics, Budapest, Hungary

Abstract
The paper summarises the main challenges and problems related to transparency in production firms and networks. The transparency problem of production are treated from three aspects, i.e., factory level data gathering tight-coupled with productions simulation; extracting knowledge from large, complex, time-dependent noisy and anomalous process logs; and identity-based tracking and tracing services within and beyond organizational borders. For all these fields of industrial relevance novel approaches and solutions are presented.

Keywords:
Transparency, Production; Information and communication technologies, Data mining, Auto identification

1 INTRODUCTION

The real-time management of changes and disturbances in production execution along with the distributed effort of supply chain planning and management have become significant in technical, economic and operational aspects of today's production companies. The practical feasibility of most of the approaches to production modeling and control boils down to providing sufficient information about the involved processes and entities [1].

In order to master the high dynamics in the processes and demand, real-time feedback from production is required. For real-time control actions, information about the state of the controlled system must be provided without large time lags. While information flow is easier to manage within and between IT components, it may become critical to maintain links between physical products and the related software agents, as the product is continually changing and moving without a permanent network connection being guaranteed [2]. Better information flow and transparency may also contribute to further improvement, such as real event-driven control [3], as well as ‘plug and produce’ performance [4] based on autonomous resources and intelligent products.

The paper addresses the transparency problem from three aspects. Firstly, data gathering on factory level is focused on, with special emphasis upon the mirroring of the current state of the production into the digital world in order to support production planning and control decisions (Section 2). Unfortunately, the huge amount of data collected during production needs data filtering and mining, because of missing and / or distorted information. Novel approaches to this issue are presented in Section 3. Section 4 concentrates on tracking of products and product data within and beyond organizational borders by introducing an open-source solution platform for this purpose.

2 DATA GATHERING AND SIMULATION AT FACTORY LEVEL

One of the utmost requirements regarding transparency in manufacturing systems at factory level is the availability of actual, reliable productions systems’ status data transformed and provided as information according to the users’ requirements.

In a large-scale manufacturing environment taking the most appropriate control decision, as well as the prediction of waiting times, workloads or utilisation of the resources is a difficult task and hardly depends on the availability of data. These control decisions and the analysis of their effects can be supported by simulation-based analysis relying on proper data acquisition and data transformation methods [5], [6].

A discrete-event simulation-based analysis system is proposed in this section which supports both the automatic mirroring of the current production system’s state in a digital model, and the analysing and validating of different production control settings and rules in the digital model.

Due to the large number of the resources, the relatively long and frequently reengineered routings and the often changing product types, the maintenance of the simulation model and the provision of up-to-date input data are always difficult tasks [7]. Thus, one of the limitations of using simulation in the on-line decision making process is the considerable amount of time spent on gathering and analysing data. In quasi real-time control (hours, minutes), however, the three key issues are.

- Data acquisition and validation for simulation input.
- Quick response time of simulation runs and analysis.
- The ability of creating the snapshot of the physical system status in the simulation model by instantaneous feedback.

The main goal of the research and development presented here was to enhance the simulation-based analysis and dispatching system by eliminating the manual data acquisition through automatic interfaces, to create a more realistic model of the factory and to improve the dispatch logic of the control system. Furthermore, the self-building simulation model now provides prospective (e.g., locating anticipated disturbances, identifying the trends of designated performance measures), and retrospective (e.g., gathering statistics on resources) simulation functionalities.

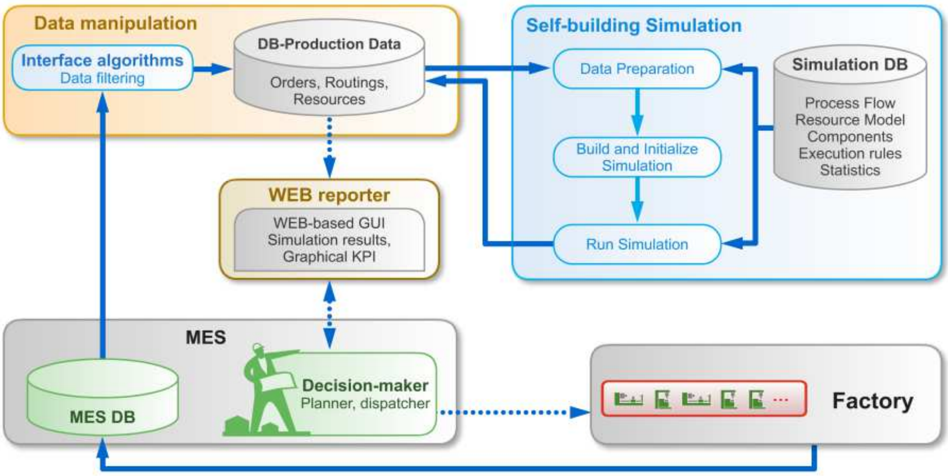


Figure 1: The structure and the dataflow of the new self-building simulation architecture

2.1 A possible architecture

Figure 1 represents the architecture and the dataflow of the developed simulation-based decision support system.

In the majority of industrial environments all the inputs of a simulation study need analysis, filtering and pre-processing for several reasons as follows:

- Both the planning side in the Enterprise Resource Planning (ERP) system with basic data and the execution side in Manufacturing Execution System (MES) may contain incorrect data entries.
- Some of the inputs (e.g., failure rates, process times, operator times, etc.) should be estimated or calculated.
- The logical and the physical organisations of the ERP and MES database are not suitable for direct use by simulation.
- Having a large-scale complex manufacturing system, both the ERP and the MES databases may include millions of records. In a case like this, specific data management techniques might be necessary to handle the data for fast coupling the digital and physical worlds.

The interface algorithms split the aggregated data of the MES into separated inputs and perform the necessary statistical analysis and process time calculations. The simulation model keeps its own, internal, simulation-specific database to support fast response time. Regarding one analysis scenario the exchange of the data is necessary only in the early stage of the simulation when the automatic model building and initialization processes are performed.

The core simulation process is depicted in the right box of Figure 1, while the detailed description of each step is given in [8]. In the last phase, the results of the simulation are inserted into the production database from where they are presented to the Decision-makers through a user-friendly web-reporter system. The necessary actions are always taken by the job-shop personnel responsible.

In case a simulation system is combined with the production database of the factory, it is possible to instantly update the parameters in the model and use the simulation parallel to the real manufacturing system supporting and/or reinforcing the decisions on the shop-floor. The system presented gives a flexible tool for engineers on the shop-floor and allows them to easily configure the simulation model of their manufacturing system and evaluate both global and local production control decisions in different circumstances, while collecting and maintaining important production related data as well.

3 TRANSPARENCY INCREASE VIA DATA MINING

In large-scale, complex and constantly changing manufacturing processes the process models and parameters used by the planning and execution systems might become out of date relatively frequently [9]. The availability of detailed event logs from various machines or tracking systems may help in automating the model refinement by using data mining and process mining algorithms. Data- and process mining allow the discovery of process routing, correlation between attributes or events, potential hidden states and operations, true task times or machine related attributes. The planning and scheduling steps can be improved by feeding back the values and model changes learnt from the event logs.

However, current process mining methods and tools, usually, cannot handle the simultaneously existing attributes of a real manufacturing process, such as longer loops (e.g., rework or re-entrant workflow), changes of mid-process product type or process evolution over time. When manufacturing execution logs are reported manually by the operators on the shop floor, the logged events may contain consequent noise (e.g., misreported machine) atop of the natural variations of event timing (e.g., completion time variability), or the interpretation of the event attributes might vary between the shop floor and the information system (e.g., the so-called ownership time duration is logged instead of the actual processing time duration). Figure 2 displays a typical scenario for differences between a real, non-overlapping batch of tasks and the logged version of the same production process.

Consequently, Figure 3 displays the ownership duration' decomposition into time segments.

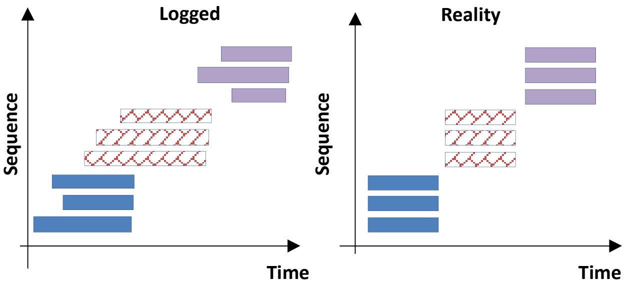


Figure 2: Logged vs. reality events Gantt chart.

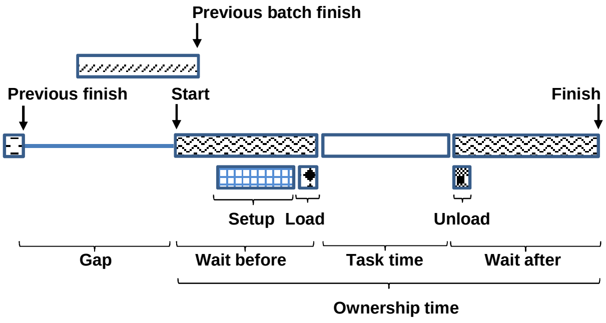


Figure 3: The ownership time duration decomposition.

Relying on an unfiltered, anomalous and inconsistent show-floor log may lead to bad estimation about the true capabilities of the factory.

The obvious resolution to the log problem is to fix the information system and have worker trainings to reduce data entry errors. However, there are several additional problems associated with this approach:

- The existing information system responsible for the tracking cannot be changed due its vendor is not available anymore.
- The changes required in the information system are costly and hard to justify.
- To fix the issues with the issue tracking, the system developers need to figure out what is wrong by using statistical and data-mining techniques anyway.
- Even when the information system is fixed, the historical data still remain incorrect.

Instead of the obvious resolution, a novel analysis framework and methodology is suggested which has the objectives to:

- discover the true properties of the manufacturing process (e.g., routings, product categories, machine capabilities, scheduling, dispatching rules and processing times) in face of the various noises,
- help answer the “What was going on in the factory at a certain time?” by applying fixes to the logs based on the more realistic properties discovered,
- provide explanations to various logging anomalies which can be used in the information system improvements,
- help evaluate and compare the key performance indicators of the manufacturing process before and after applying the fixes.

In order to find the non-stationary routing graph (or net) and time-related properties of various product types from noisy, anomalous operation logs, an iterative framework is proposed. In the framework partially discovered properties are used for reasoning about the effects present in the logs and confirming or abandoning assumptions about the sources of the anomalies. The analysis framework requires some basic assumptions to be met in the underlying manufacturing process:

- only one operation should be performed on the product at a time in a task (e.g., no parallelism or task overlapping present),
- a machine performs a single operation on a (set of) products at a time.

In general the log files might lack important information (e.g., task start time) or the reasons of certain log anomalies may be completely different in other cases, rendering the corrective measures inert or even corruptive. A different case might require constant and extensive help from an expert of the manufacturer and expect to revalidate and rethink the existing assumptions. The presented framework may be used as a strong base for developing concrete, tailored applications for these new and different cases.

The diagram in Figure 4 shows the log analysis and knowledge discovery steps of the Analysis Framework.

The methodology of the knowledge discovery process can be summarized as follows:

- discover the high-level structures at first (such as the routing graph, product categories and machine groups) which ought to have large amount of samples,
- discover the lower-level structures, (such as time-related properties, dispatching rules, scheduling, etc.) with the guidance of the already discovered high-level structures,
- overcome the various noises in the log by iterating until the consistency of the discovered properties is achieved,
- fix the noises between iterations and between the high-level to lower-level discovery steps and
- if the sample amount for a particular case is too low for a consistent discovery, try using the implicit knowledge of the production environment (such as when a lunch break usually happens, what were the other tasks at the time, etc.)

The extracted finer knowledge may help in several key areas of the manufacturing:

- highlights the misbehaviour of the shop-floor personnel in terms of different logging anomalies,
- allows the scheduling system to use a more detailed routing information which includes details of product type changes and reworks,
- allows better what-if simulations to be carried out and provides additional information for factory optimizations,
- makes the automatic maintenance of the manufacturing attributes possible in various shop-floor systems and
- increases the overall transparency of manufacturing and shop-floor activities.

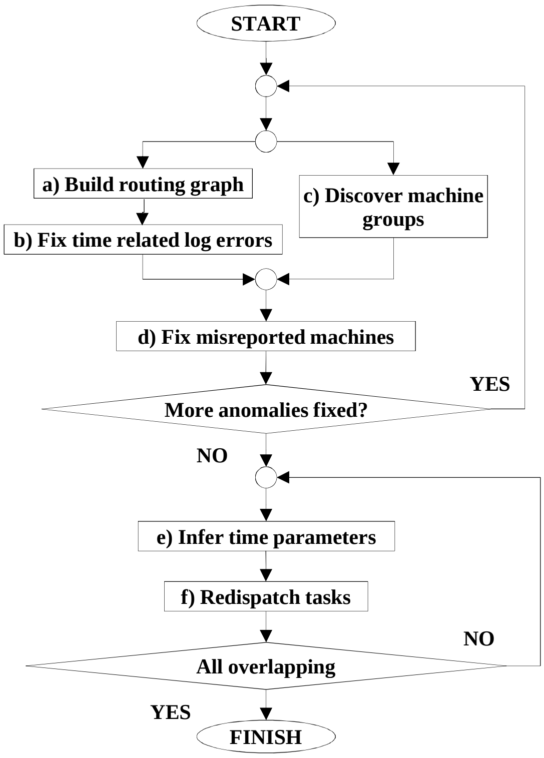


Figure 4: Workflow of the analysis framework.

4 IDENTITY-BASED TRACKING AND TRACING OF PRODUCTS AND PRODUCT DATA

Recent trends in industrial production are marked by a growing demand for improved observability in production, delivery, usage and disposal of products. This is, in numerous cases, owing to product customization gaining a foothold even in mass production, customers beginning to demand improvement in up-to-date information about the products they ordered. Also, the efficient handling of quality feedback is becoming a decisive market advantage (including, but not restricted to, product recall campaigns if faults are discovered). In addition to changing demands, the typical structure of industrial production is undergoing changes: more and more often, companies are not involved in a fixed *supply-chain*; instead, they take part in *production networks* where, in some cases, participants team up only for the procurement of a single product and act as competitors in other cases. This triggers the need for measures which can securely deliver timely and error-free information about material movement in production; often, on the level of individual work-pieces or items.

Tracking is the term referring to the act of observing, in most cases, the spatial motion of an entity. This may be implemented in the form of *checkpoint and time-stamp* pairs (i.e., the presence at discrete, pre-defined points or stages is checked), or *spatial coordinates* (GPS data, length of path travelled etc.) may be used instead of the poor resolution of discrete points. Tracking is mainly used for providing information about the advance of a given shipment (e.g., to meet correction measures in case of lagging delivery), while aggregated tracking data may be used for optimization of future logistics operations.

Tracing, on the other hand, is less concerned about the movement of the item alone. Instead, it focuses on changing relations with the production environment, including other components the given entity may enter a relation with. Such cases occur in the assembly of a composite product from sub-components or the mixing of materials (e.g., food and chemicals). Tracking data gathered over a manufacturing process result in a so-called *gozinto graph*, expressing assembly relations by describing which components *go into* (become parts of) other ones. The most straightforward use of tracing is quality control, ranging from directed product recall campaigns to improvement of manufacturing technology or work processes relying on aggregated data.

Nowadays, the majority of applications which keep track of items throughout significant segments of their life cycle rely on *tracking-based operations*. However, if relevant parts of the life cycle take place under the authority of other parties or the complexity of the value creation process (e.g., multiple suppliers or customers) require maintaining tracking or tracing across organizational borders, *advanced item-centric services* become necessary. In such cases, item-related data and services (e.g., notification, subscription) are shared among process participants with proper access restrictions.

4.1 An open-source solution platform for tracking and tracing

The primary goal of the European research project “TraSer” (Identity-Based Tracking and Web-Services for SMEs) was the development of a free, open-source software package offering affordable item-centric tracking and tracing solutions for small and medium-size enterprises (SMEs) operating in environments transcending company or organizational borders. A special release of the solution platform was adapted to tracking product data files (e.g., blueprints produced by several industrial partners during collaborative product design). Data tracking has particular relevance as today, company assets more and more dominantly exist as electronic data that have to be consistently guided through their intended life cycle and protected from unauthorized access.

The TraSer solution platform is based on communication through web services (WS), a standard which is by now well-established and supported by numerous commercial and free frameworks. Web services also allow data exchange to be flexible and easier to configure than most of the “conventional” EDI (electronic data interchange) channels. The TraSer solution platform allows participants to build a TraSer network where two components can be distinguished: nodes (servers) and clients. (See also Figure 5 with a simple TraSer network.)

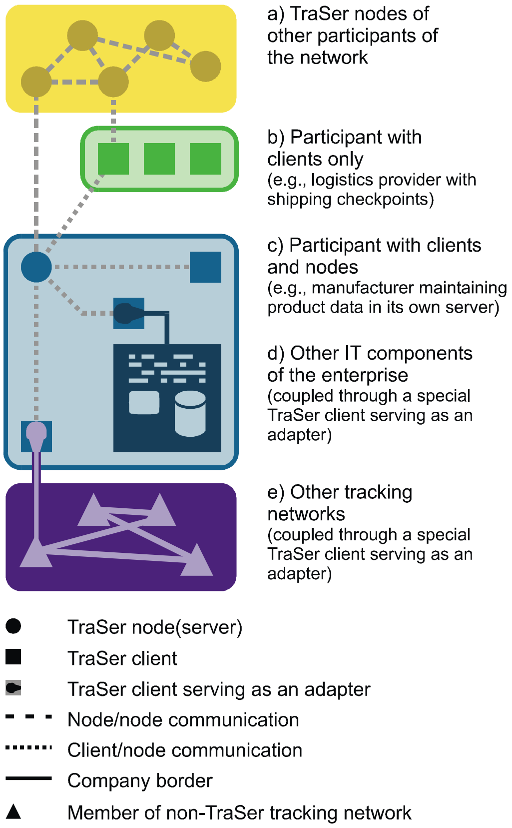


Figure 5: Simplified example of a TraSer network [10].

TraSer nodes are the servers responsible for storing item-related data and providing authorized parties with access to them. The nodes receive requests in the form of XML queries which allow more flexibility in customizing the data models used by the partners. TraSer nodes can also forward queries or updates to each other—these node-to-node connections span a part of the TraSer network (See Figure 1, part a) depicts a small network of several nodes and part c) stands for a company which operates a TraSer node within its own IT infrastructure).

The forwarding of queries or updates is required if parts of the item information in question are maintained by different nodes. This practice is typical for production networks where the products of several manufacturers are combined to a composite item—in that case, information about the sub-assemblies possibly resides in different TraSer nodes of the same TraSer network. Also, the data of a given item can be extended by further properties which are not necessarily located in the same node (e.g., when a manufacturer wishes to add its own relevant notes to the item description of its supplier). Therefore, TraSer nodes can forward queries or update requests to each other.

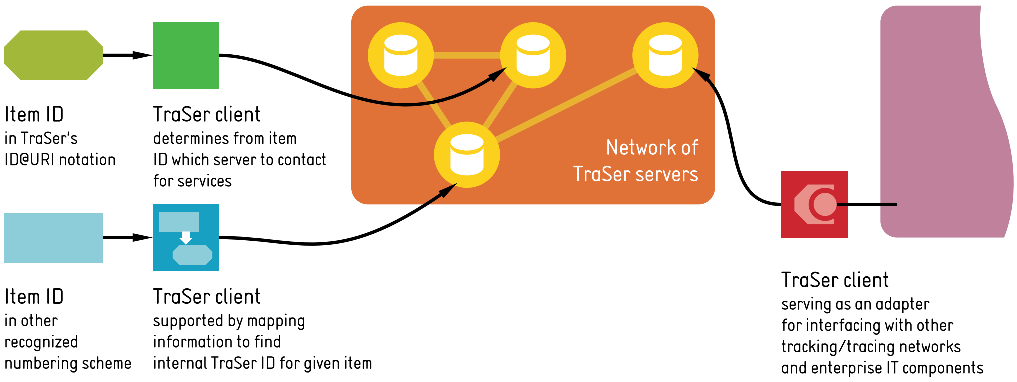


Figure 6: Access to item-related services in a TraSer network using various methods of unique identification and specialized clients using different interfaces. Freedom in addition of customized clients allows the adaptation and coupling of a TraSer-based solution with other components and systems

Adding a new node to the TraSer network (e.g., when a new manufacturer joins the community) consists in matching the data models and messages to the requirements of the network and setting the access rights for existing network members if required.

TraSer clients form the other main group of components in a TraSer network. Clients connect the nodes with the rest of the world by providing external interfaces and addressing one or more nodes with item-related queries or updates. Clients can be fitted with various kinds of interfaces, i.e., they can be designed for human operators, peripheral devices (readers etc.), other components of the enterprise infrastructure (stock management, ERP), and other tracking and tracing systems. Since TraSer interface specifications are freely available, users can develop specific clients tailored to their given needs.

Figure 5 shows several specialized cases of client use. Companies which do not have their own TraSer-tracked items (e.g., logistics partners not maintaining their own transportation asset data in a TraSer system but updating the shipment information of a manufacturer's products; or a small supplier which lets a larger partner care about hosting its product data) are not required to operate their own node, as shown in part b). The company in part c) operates, aside from the TraSer node, several specialized clients. One of these serves as an adapter for accessing other components of the manufacturer's IT infrastructure (part d), while another client was customized as an interface towards another tracking network (part e).

If consistency of transactions and protection from eavesdropping or intrusion by unauthorized parties are a concern, several implementations of the relevant web service standards (WS-Reliable Messaging, WS-Security) can be used.

While developing and elaborating product ontology is hardly within the scope of the TraSer project, an actual instance of a TraSer network can rely on several previous achievements in the field. Within the project, possibilities of compliance with the UN/CEFACT Core Component Library are examined, while the integration of other sources (such as the interoperability-related network of excellence INTEROP [11] and related projects [12]).

TraSer relies on web services for communication. Item-related data are maintained by TraSer servers which communicate with each other if queries or updates are forwarded or broken down to distributed components. TraSer clients serve as interfaces "to the rest of the world", including human operators, automated checkpoints or other components of the given IT infrastructure (see also Figure 6).

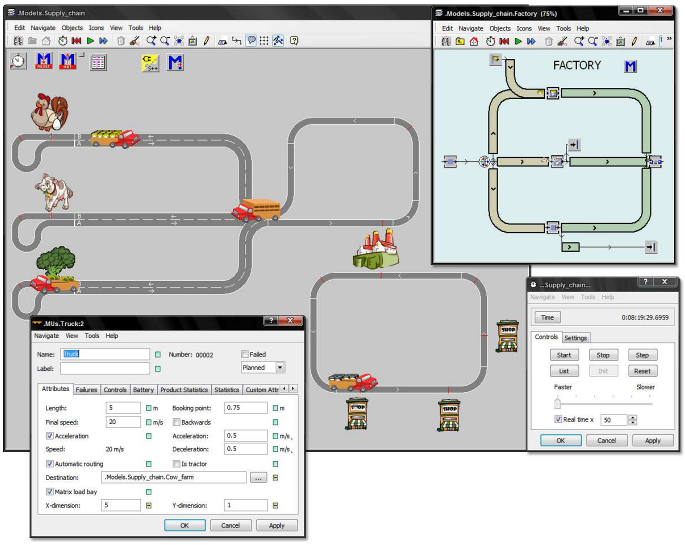
4.2 Pilot applications

Asset identification. Here, the goal is to identify vehicles and roll cages entering or leaving the facilities of a postal delivery firm. Suitable equipment was, in part, already installed in a given logistics center, and TraSer nodes can readily take over the task of enter/leave transactions. The new solution based on TraSer offers a good basis for further enhancement as well: specific clients could report yard traffic to office personnel or other components of the enterprise IT infrastructure.

Asset tracking. Recording and forwarding vehicle movement, as described above, is, in fact, already leading to a higher functionality level, as it introduces tracking services. This phase allows the transparent surveillance of departure or arrival at several client-equipped locations, enabling the progress of logistics processes to be monitored. The extension of tracking to smaller transportation assets (specifically roll cages) makes it possible to give specific instructions for loading and unloading vehicles, checking the contents of a vehicle, and keeping track of the location of roll cages, thus helping to prevent loss, theft, shortage or surplus build-up of roll cages at various locations. Applying mobile clients would, in this stage, provide a cost-efficient alternative to installation of TraSer clients at destinations which are less frequented by Itella's deliveries.

Asset-based tracking of goods. Here, the transportation assets tracked by the TraSer network are still in a closed circulation (as are their IDs), however, the goods moving together with the roll cages are usually participating in a flow-through supply chain and the latter items only appear in Itella's tracking system once, for a limited amount of time or logistics operations. These goods belong to companies using Itella's logistics services, and Itella can, by providing them with information about delivery progress, offer them goods tracking services, either through a human-readable web interface, or, in more advanced cases, through giving them limited access to the TraSer network by specific TraSer clients.

Tracking of perishable product in a food supply chain. Ambient conditions during transportation can affect the quality of the products being delivered—this holds all the more for food supply chains. In addition to installing appropriate appliances for keeping transportation conditions (e.g., temperature, humidity) within tolerances, the conditions are more and more often regularly checked and reported both as a credential measure for guaranteed quality, as well as for limiting responsibility in case of quality complaints. Transparent reporting is demonstrated in our example by simulated temperature readings unique to each vehicle taking part in the logistics processes.



Name: Truck
Label:
Number: 00002
☐ Failed
Planned
Length: 5 m
Booking point: 0.75 m
Final speed: 20 m/s
☐ Backwards
☑ Acceleration
Acceleration: 0.5 m/s
Speed: 20 m/s
Deceleration: 0.5 m/s
☑ Automatic routing
☐ Is tractor
Destination: .Models.Supply_chain.Cow_farm
☑ Matrix load bay
X-dimension: 5
Y-dimension: 1

Time 0:08:19:29.6959
☑ Real time x 50

Figure 7: Application of the TraSer platform—tracking in a simplified food supply chain

Measurement occurs in certain time intervals and the values are registered in the history of the given vehicle. Since the goods transported in the vehicle are logically linked to the observed transportation asset, the values of temperature readings can be automatically forwarded to the corresponding attributes of the transported products where they permanently remain to allow subsequent quality checks and tracing of an item's history (Figure 7).

5. Conclusions

The paper addresses the transparency problem of production from three aspects, i.e., factory level data gathering tight-coupled with productions simulation; extracting knowledge from large, complex, time-dependent noisy and anomalous process logs; and identity-based tracking and tracing services within and beyond organizational borders. All these fields are of industrial relevance, moreover, the introduced solutions have been installed in industry, at least as pilot cases.

Transparency in production and supply chain is certainly a field, where the evolution of the information and communication technologies (ICT) and manufacturing will probably proceed hand in hand: the former can receive real challenges from the latter, which, in turn, will have more and more benefits in applying novel ICT solutions

Acknowledgements

The authors acknowledge the support of the Hungarian Scientific Research Fund (OTKA) through grant T-73376 "Production structures as complex adaptive systems", and of the National Office for Research and Technology, Hungary (NKTH) through grant "Digital, real-time enterprises and networks" OMFB-01638/2009. A part of the activities described here was financed by the EU-project "Development of tools to communicate advanced technologies on active and intelligent packaging".

5 REFERENCES

[1] Jansen-Vullers, M.H., van Dorp, C.A., Beulens, A.J.M., 2003, Managing traceability information in manufacture, Int. J. of Inf. Managm., 23/5, 395–413.

[2] McFarlane, D., Sarma, S., Chirn, J.L., Wong, C.Y., Ashton, K., 2003, Auto ID systems and intelligent manufacturing control, Eng. Appl. of Artificial Intelligence, 16/4, 365-376.

[3] Schuh, G., 2006, Sm@rt logistics: Intelligent networked systems, CIRP Annals, 55/1, 505-508.

[4] Scholz-Reiter, B., Freitag, M., 2007, Autonomous processes in assembly systems, CIRP Annals, 56/2, 712-729.

[5] Kádár, B., Lengyel, A., Monostori, L., Suginishi, Y., Pfeiffer, A., Nonaka, Y., 2010, Enhanced control of complex production structures by tight coupling of the digital and the physical worlds, CIRP Annals, 59/1, 437–440.

[6] Bagchi, S., Chen-Ritzo, C., Shikalgar, S.T., Toner, M., 2008, A full-factory simulator as a daily decision-support tool for 300mm wafer fabrication productivity, Proc. of the 2008 Winter Sim. Conf., 2021-2029.

[7] Pfeiffer, A., 2007. Novel methods for decision support in production planning and control, PhD Thesis, Budapest University of Technology and Economics.

[8] Kádár, B., Pfeiffer, A., Monostori, L., Popovics, G., Vén, Z., 2011, Co-analysing situations and production control rules in a large-scale manufacturing environment, Proc. of the 44th CIRP Conf. on Manuf. Systems, June 3-5, Madison, WI, USA (in print)

[9] Karnok, D., Monostori, L., 2010, Routing model refinement in large scale manufacturing environment by using data mining. Proc. the of 43rd CIRP Conf. on Manuf. Syst., May 26 - 28, Vienna, Austria, 725-734

[10] Monostori, L., Kemény, Zs., Ilie-Zudor, E., Szathmári, M.,, Karnok, D., 2009, Increased transparency within and beyond organizational borders by novel identifier-based services for enterprises of different size, CIRP Annals, 58/11, 417-420.

[11] Panetto, H., Scannapieco, M., Zelm, M., 2004, INTEROP NoE: Interoperability research for networked enterprises applications and software, Lecture Notes in Computer Science, 3292, 866–882.

[12] Berre, A.J., Elvesæter, B., Figay, N., Guglielmina, C., Johnsen, S.G., Karlsen, D., Knothe, T., Lippe, S., 2007, The ATHENA interoperability framework, enterprise interoperability II, Proc. of I-ESA 2007, March 27-30, Madeira, Portugal, 569–580.