SILESIAN UNIVERSITY OF TECHNOLOGY PUBLISHING HOUSE

SCIENTIFIC PAPERS OF SILESIAN UNIVERSITY OF TECHNOLOGY 2024
ORGANIZATION AND MANAGEMENT SERIES NO. 212

PROCESS MINING OF PRODUCTION SUPPORTED BY MES SYSTEMS – A CASE STUDY

Piotr JANKE[1*], Grzegorz ZIELIŃSKI[2]

[1] Silesian University of Technology, Faculty of Organization and Management; piotr.janke@polsl.pl, ORCID: 0000-0001-8065-9013
[2] Muller - Die lila Logistik Polska; Grzegorz.Zielinski@lila-logistik.com
* Correspondence author

Purpose: The primary purpose of this paper is to use a process mining method on collected and appropriately prepared event logs from the MES class system of the entity under study.
Design/methodology/approach: With the classical process mining approach using the XES standard, the Alpha algorithm and Petri nets, event logs of the manufacturing process were analysed. To perform the analysis, 379,341 rows of data were extracted from the MES system, representing 7194 instances and 76 different activities.
Findings: a key result of the presented research is the conclusion regarding the quality of the data generated by the controllers. Timestamps for events in the MES system may need to be corrected.
Research limitations/implications: The paper presents a case study for a specific set of controllers and an MES system. For this particular configuration, additional mechanisms are required to synchronise time stamps from one location (e.g. an NTP server). The results may be different for a different hardware and software configuration of the MES system.
Practical implications: The result of the research carried out in this paper also leads to conclusions of a practical nature. Incorrect synchronisation of times makes it impossible to control and analyse the process as a whole, even when the whole process is working correctly as an assembly of individual operations. The guidelines developed were passed on to those responsible for the operation of the MES system.
Originality/value: In the literature, various case studies can be found for process mining including event logs of ERP systems of others. The authors did not find similar studies for MES systems. In particular, the classical approach for log preparation requires awareness of potential time synchronisation problems.
Keywords: manufacturing execution system, process mining, plc.
Category of the paper: Research paper.

1. Introduction

With the dynamic development of technology, the evolution of manufacturing processes has become an integral part of modern businesses. Based on Michael Hammer's literature, it can be concluded that the evolution of manufacturing processes is essential for maintaining a competitive advantage in today's dynamic market. Companies that continually improve their processes can respond quickly to changing customer needs and market trends, resulting in greater customer satisfaction and a greater chance of success (Hammer, Champy, 2009). In response to these challenges, exploring manufacturing processes has become an important area of research.

Process mining is an emerging discipline that is based on process models and data mining. Production process mining refers to the analysis and study of activities undertaken to identify and evaluate new opportunities for improvement and optimise existing processes in manufacturing facilities. It aims to increase productivity, improve quality, reduce costs, minimise cycle time and make production more flexible and reactive.

At present, the use of MES-class information systems to support processes at the activity level in production operations is becoming increasingly popular in many manufacturing companies. The use of modern methods to analyse processes at the level of individual production stages appears to be an interesting area of scientific interest.

The main objective of the paper is to analyse the production process in a company using process mining methods on collected and appropriately prepared event logs from the MES class system for a selected production process of the studied entity.

2. Theoretical Background

The classic approach to organising a manufacturing process in ERP (or MRP) systems involves defining articles, structures (BOM) and processes for the structures. The tasks in the process should have defined resources. It is only by linking the resources to a database using controllers that keep the states of these resources up-to-date that the production process can be precisely controlled.

2.1. Manufacturing Execution Systems

MES (Manufacturing Execution Systems) play a key role in modern production management. MES are integrated IT systems that monitor, control and optimise production processes in real time. They make it possible to track every stage of production, from raw materials to finished products, thereby increasing efficiency, reducing costs and improving quality.

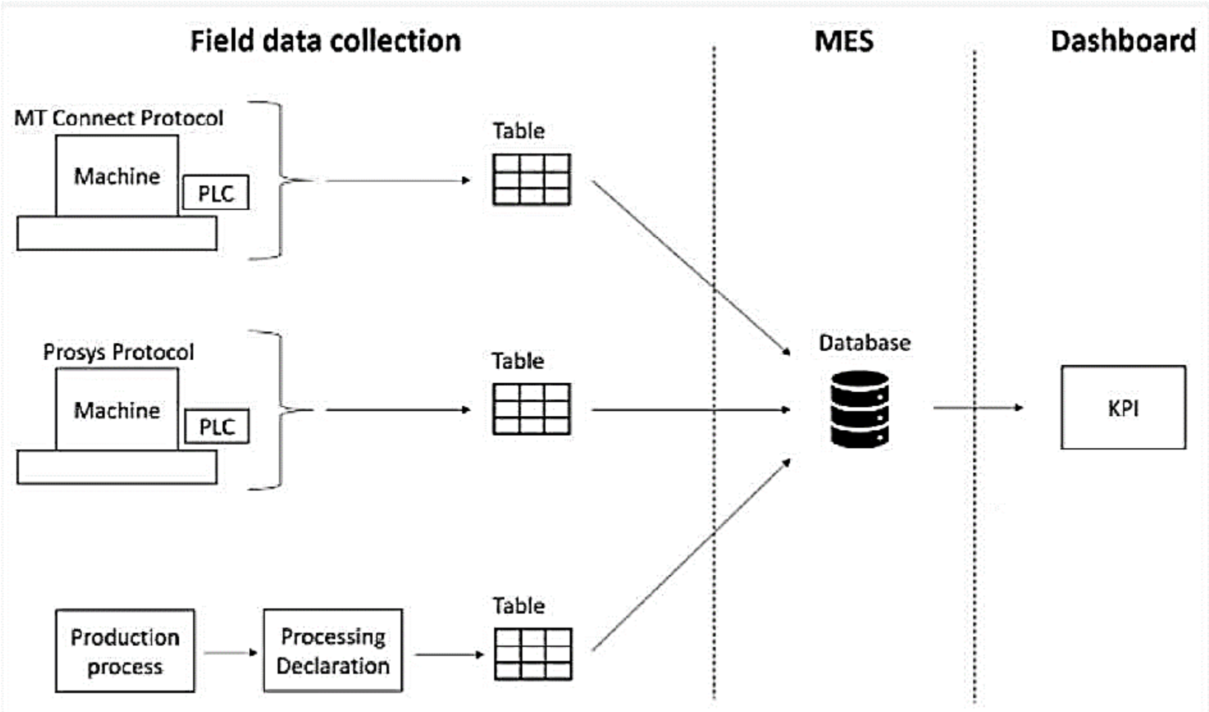


Figure 1. Machine data collection in the MES system.

Source: (Bianchini et al., 2024).

One of the main tasks of MES is to collect and analyse production data. These systems integrate with other management systems, such as ERP (Enterprise Resource Planning), enabling full synchronisation of data and processes across the enterprise. For example, MES can automatically update stock levels based on actual raw material consumption, which minimises the risk of production downtime.

The literature highlights the numerous benefits of implementing MES, such as increased transparency in production processes, better resource management and the ability to respond quickly to changes in market demand (Johnson, 2021; Smith, 2020). In addition, MES support Industry 4.0 initiatives by enabling integration with IoT (Internet of Things) technologies and cloud-based data analytics (Brown, 2022).

2.2. Process Mining

Process mining technology is gaining importance as a tool for analysing and optimising business processes. In recent years, research in this area has focused on integrating with modern technologies such as artificial intelligence (AI) and machine learning (ML) for more advanced process data analysis. One key area of research is the use of AI and ML to automatically discover process models from event log data. Research shows that these technologies can significantly improve the accuracy and efficiency of process analysis (Wil van der Aalst, 2024). Furthermore, these techniques enable the prediction of future events and the identification of potential problems, which is crucial for proactive process management (Attias, 2024). Another important line of research is the application of process mining in various industry sectors.

Examples include applications in healthcare, where the technique helps to optimise patient pathways and improve the quality of care (Rebuge, Ferreira, 2012). In the manufacturing sector, process mining is used to identify bottlenecks and optimise production flows (Thiede et al., 2018).

Process mining combines traditional model-based process analysis and data-driven analysis techniques. As shown in Figure 2, process mining can be seen as a link between data science and process science. Process mining seeks a confrontation between event data (i.e. observed behaviour) and process models (either manually performed or automatically detected).

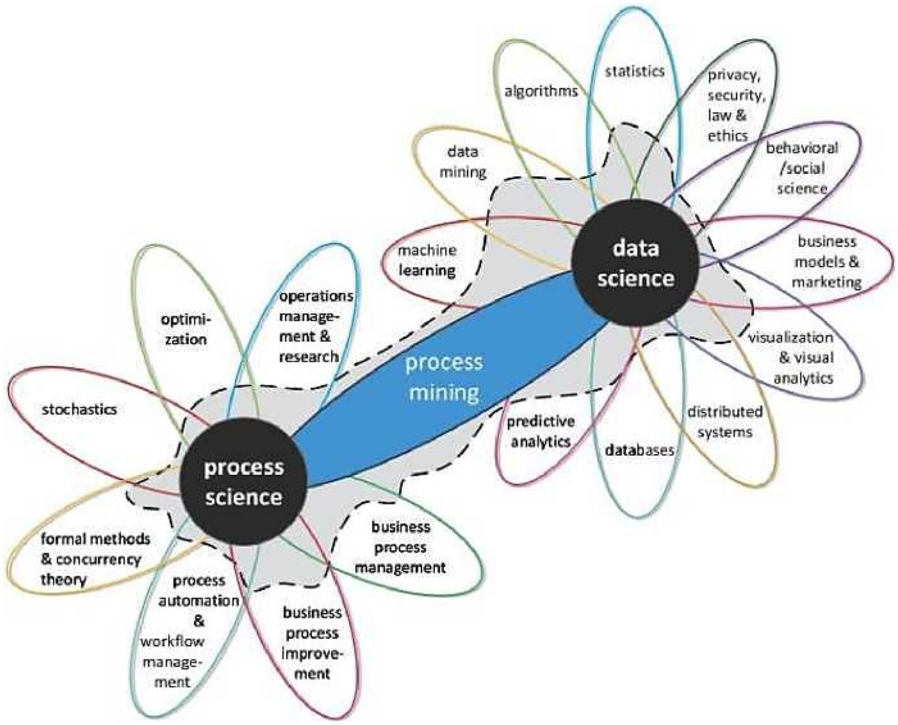


Figure 2. Illustration of the link between process science and data-driven techniques.

Source: van der Aalst, 2016, p. 18.

The main approaches to data science tend to be process-independent. Data mining, statistics and machine learning techniques do not consider comprehensive process models. Process science approaches are process-oriented, but often focus on modelling rather than learning from event data. The unique positioning of process mining makes it a powerful tool for using the increasing availability of data to improve complex processes. Process mining has only recently emerged as a sub-discipline of both data science and process science, but the relevant techniques can be applied to any type of operational process (organisations and systems) (van der Aalst, 2016).

3. Methods

On the methodological side, the process mining technique involves the use of several tools and standards implemented in the explorer. In this paper, the ProM environment version 6.12 was used, as well as a standard sequence of steps for the prepared event log including data conversion to the XES standard and then the use of the Alpha algorithm allowing for the preparation of a process model in the form of a Petri net. The result enables the use of further plug-ins such as conversions to the BPMN standard diagram, among others.

XES standard

The XES standard defines a grammar for a markup-based language, the purpose of which is to provide information systems designers with a standardised and extensible methodology for capturing the behaviour of systems using the event logs and event streams that are defined in the XES standard. This standard includes an XML schema describing the structure of an XES event log/stream and an XML schema describing the structure of an extension of such a log/stream. In addition, this standard contains a core set of so-called XES extension prototypes that provide semantics for certain attributes recorded in the event log/stream ('IEEE Standard for EXtensible Event Stream (XES) for Achieving Interoperability in Event Logs and Event Streams,' 2023)

Alpha algorithm

The α or α -miner algorithm is an algorithm used in process mining to reconstruct causality from a set of event sequences. It was first proposed by van der Aalst, Weijters and Maruster. The aim of Alpha Miner is to transform an event log into a workflow network based on the relationships between different activities in the event log. An event log is many sets of traces, and a trace is a sequence of activity names (van der Aalst et al., 2004).

Alpha Miner was the first process discovery algorithm ever proposed and gives a good overview of the purpose of process discovery and how different actions are performed within a process. Alpha miner was also the basis for the development of many other process mining techniques, such as heuristics and genetic exploration.

The Alpha algorithm starts by transforming the event log into direct trace, sequence, parallelism and selection relations and then using these to create a Petri net describing the process model. Initially, the algorithm constructs a trace matrix. Using the trace matrix, a process model can be constructed. Based on the relationships, a trace-based matrix is discovered. Using the trace-based matrix, locations are discovered.

Petri nets

Petri nets are the oldest and best-studied process modelling language for modelling concurrency. Although the graphical notation is intuitive and simple, Petri nets are executable and many analytical techniques can be used to analyse them (van der Aalst, 2016).

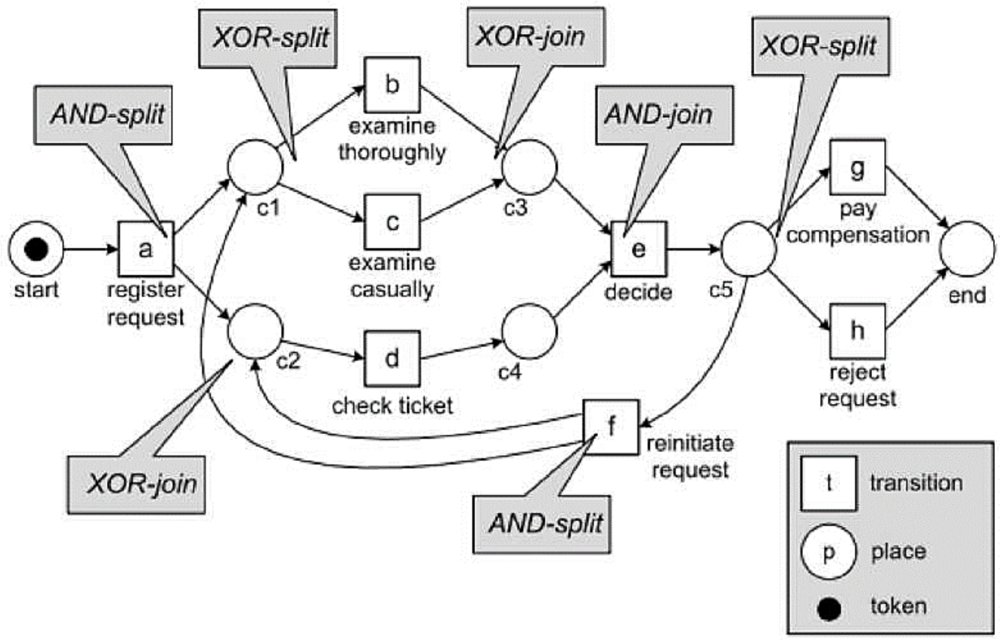


Figure 3. Example of a Petri net.

Source: van der Aalst, 2016, p. 60.

Figure 2 shows an example Petri net with labelled logic gates, locations and transitions. The AND and XOR gates represent transitions to the next place in the process or to the next action.

BPMN notation

Business Process Modelling Notation (BPMN) has become one of the most widely used languages for modelling business processes. BPMN is supported by many tool vendors and has been standardised by OMG. Figure 3 shows a process model in BPMN notation.

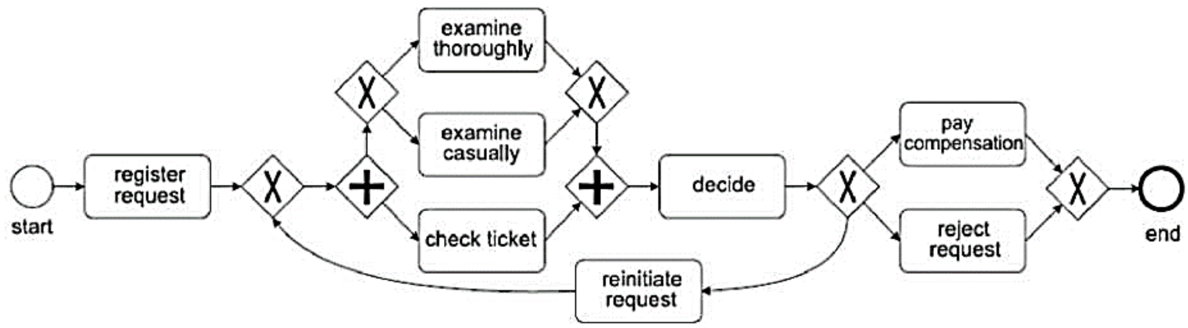


Figure 4. Process model in BPMN notation.

Source: van der Aalst, 2016, p. 69.

Business Process Modelling Notation (BPMN) is a graphical notation that represents the stages of a business process. BPMN represents the end-to-end flow of a business process. The notation is specifically designed to coordinate the sequence of processes and messages that flow between different process participants in a related set of activities (Facility Management Group…).

3.1. Data sets

Each action that has been performed by an MES system command has been recorded in the system, a time stamp has been given and, depending on the type of task (device type), the data from the process performed, the data collected by the scanner or the logical value of the sensor. In the system, the collected data is assigned to a particular device, which can perform one or more operations on one or more programmes.

The raw data extracted from the MES system is not suitable for mining, it must first be properly prepared. Two types of files have been generated from the system from which the data will be extracted. Their adaptation to a common form differs. In the files that are generated from the process equipment, all blank rows must first be removed. In the next step, a description of the activity that was performed was added to each record. In the last step, redundant data has been deleted.

The second type of data that was extracted from the system is the data that the scanner collected. Once exported, all that was needed was to remove the redundant columns.

For the analysis, 379,341 rows of data were downloaded from the MES system, representing 7194 cases and 76 different activities. The downloaded data were reduced to a common form, which is shown in Table 1.

Table 1.
Single case of analysed data

JOB	ID	DATA
113232620545000	beam	2023-06-27 10:59:17
113232620545000	RGTP30	2023-06-27 11:04:36
113232620545000	Disc & Drum RGT	2023-06-27 11:04:36
113232620545000	Brake RGT	2023-06-27 11:04:36
113232620545000	LFTP10	2023-06-27 11:05:27
113232620545000	Disc & Drum LFT	2023-06-27 11:05:27
113232620545000	Brake LFT	2023-06-27 11:05:27
113232620545000	BrkAsm_L	2023-06-27 11:08:55
113232620545000	BrkAsm_L	2023-06-27 11:09:00
113232620545000	BrkAsm_L	2023-06-27 11:09:05
113232620545000	BrAsmRb_L_ALL	2023-06-27 11:10:27
113232620545000	RSARGT	2023-06-27 11:10:28
113232620545000	BrkAsm_R	2023-06-27 11:10:29
113232620545000	BrkAsm_R	2023-06-27 11:10:33
113232620545000	BrAsmRb_L_ALL	2023-06-27 11:10:34
113232620545000	BrkAsm_R	2023-06-27 11:10:40
113232620545000	NutRw_L	2023-06-27 11:11:01
113232620545000	RSALFT	2023-06-27 11:11:37
113232620545000	NutRw_R	2023-06-27 11:11:52
113232620545000	BrAsmRb_R_ALL	2023-06-27 11:12:06
113232620545000	BrAsmRb_R_ALL	2023-06-27 11:12:15
113232620545000	ShockAbsorberRH	2023-06-27 11:12:39
113232620545000	SqzNtLRGT	2023-06-27 11:13:06
113232620545000	ABS_L	2023-06-27 11:13:06
113232620545000	StdBlt_L	2023-06-27 11:13:19
113232620545000	StdBlt_L	2023-06-27 11:13:25
113232620545000	KAPSEL_PRAW_NISK	2023-06-27 11:13:34

Cont. table 1.

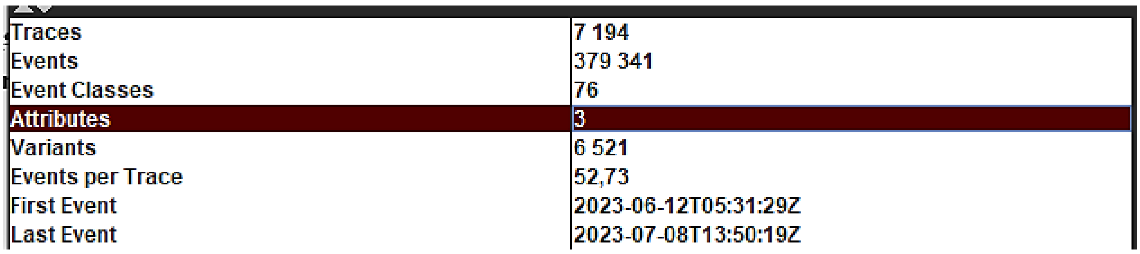
113232620545000	StdBlt_R	2023-06-27 11:13:37
113232620545000	StdBlt_R	2023-06-27 11:13:42
113232620545000	ShockAbsorberLH	2023-06-27 11:13:46
113232620545000	SqzNtLLFT	2023-06-27 11:14:17
113232620545000	ABS_R	2023-06-27 11:14:17
113232620545000	KAPSEL_LEWY_NISK	2023-06-27 11:14:45
113232620545000	BBHRGT	2023-06-27 11:15:10
113232620545000	BHRGT	2023-06-27 11:15:23
113232620545000	BBHLFT	2023-06-27 11:16:00
113232620545000	RA LFT	2023-06-27 11:16:09
113232620545000	BHLFT	2023-06-27 11:16:14
113232620545000	SCN_DISK_LFT	2023-06-27 11:16:23
113232620545000	RA RGT	2023-06-27 11:17:10
113232620545000	CLPRRGTABV	2023-06-27 11:17:10
113232620545000	CLPRRGTBLW	2023-06-27 11:17:10
113232620545000	RA__RIGHT_CHECK	2023-06-27 11:17:33
113232620545000	_Caliper_RGT	2023-06-27 11:17:37
113232620545000	CLPRLFTABV	2023-06-27 11:18:01
113232620545000	CLPRLFTBLW	2023-06-27 11:18:01
113232620545000	RA__LEFT_CHECK	2023-06-27 11:18:39
113232620545000	_Caliper_LFT	2023-06-27 11:18:40
113232620545000	P1313_L	2023-06-27 11:21:56
113232620545000	P1313_R	2023-06-27 11:21:56
113232620545000	Rack Position Axle	2023-06-27 11:22:35
113232620545000	Rack ID Axle	2023-06-27 11:22:36

Source: own work

In the table, the 'ID' column is the case number, the 'JOB' column is the classifier and the 'date' column is the timestamp for each case.

4. Results

By examining the figures from the variant flow diversity tracking module in the system, which are shown in Figure 4, it can be seen that there is a large number of different process runs in the analysed period. The table indicates a value of 6521 different runs in 7194 processes, which represents more than 90% of all results.



Traces	7 194
Events	379 341
Event Classes	76
Attributes	3
Variants	6 521
Events per Trace	52,73
First Event	2023-06-12T05:31:29Z
Last Event	2023-07-08T13:50:19Z

Figure 5. Figures from the variant tracking module.

Source: own work.

Considering the technological process of building modules and the activities where repair activities may occur the number of different flows should not be so high. Comparing the flow processes in the system to the physical process, it can be seen that the activities do not follow the technological sequence for the module build.

Figure 5 shows the part of the process from the variant that obtained the highest repeatability.

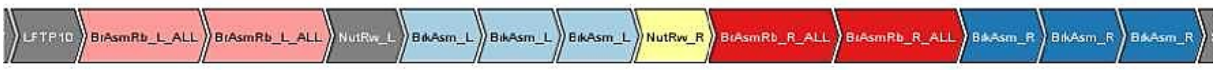


Figure 6. Most recurrent option.

Source: own work.

When investigating the problem associated with the lack of correct process sequence in the modules, the focus was on understanding exactly why this anomaly occurs. To do this, the focus was on Atlas Copco's Power Focus 6000 EC key controllers, which are responsible for data with incorrect time stamps. An analysis of the controller settings showed that there were problems with the wrong time setting on the controller. In order to set the correct data, the automatic time setting option was selected in communication with the internal NTP server.

Analysis of the results on subsequent days showed that the results continued to be marked with the wrong timestamps. After further testing to apply the correct settings, it was found that there are two main problems. Firstly, there is a problem with internally keeping the correct time on the controller. This could be due to inaccuracy of the internal clock or problems with the time synchronisation mechanism on the controller itself. Secondly, limitations have been encountered regarding the ability to increase the frequency of synchronisation with the internal NTP (Network Time Protocol) server. This means that the controller cannot update its time more frequently from the NTP server, which affects data quality.

The process between 29 and 30 June 2023 was used to generate the Petri nets. Three separate networks can be identified in the generated data. Figure 4 shows the identified network sections in the process analysed.

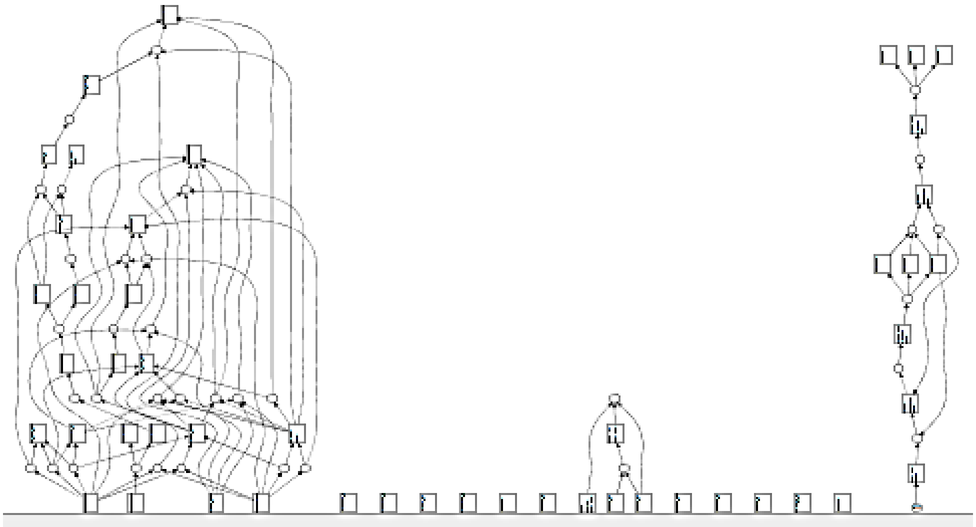


Figure 7. Petri nets generated (slice).

Source: own work.

The generated Petri nets should represent a single coherent process. The appearance of three separate, unconnected threads indicates anomalies in the data, which consist of incorrect timestamps

5. Discussion

In order to properly explore production processes, a prerequisite is to have the correct data. If one of the controllers shows incorrect hours and there is no way to change this, it is worth finding another source of data. The solution to this problem can be to give a stamp when accepting data into the host system (MES). The proposed method ensures that the data in the system will always have data from one source, which is guaranteed to be correct.

Table 2.
Proposal to add a new column for data collection

JOB	ID	DATA - controller data - transmitted by the controller	Data - data sent by the MES master system
113232631153000	ABS_L	2023-06-29 00:13:44	2023-06-29 00:17:45
113232631153000	ABS_R	2023-06-29 00:16:02	2023-06-29 00:16:57

Source: own work.

Table 2 shows how the new data should be located and where it should be taken from. Implementation should be carried out by changing the data source in the reports to that indicated.

6. Conclusions

The results of the study unequivocally confirm that it is possible to analyse and evaluate production processes on the basis of event logs from an MES-class IT production system and process mining methods. Although the investigated event logs did not contain consistent data, they made it possible to learn about the assembly process and to find an error in the system in the form of malfunctioning system clocks.

The results of the study, together with a proposal for implementing the changes, will be presented to those responsible for developing the MES system. Once the changes have been completed, the data should be collected again and the assembly process analysed again.

References

1. Attias, S. (2024). Predictive Capabilities in Process Mining. *International Journal of Business Process Management*, *8*, 78-89.
2. Bianchini, A., Savini, I., Andreoni, A., Morolli, M., Solfrini, V. (2024). Manufacturing Execution System Application within Manufacturing Small–Medium Enterprises towards Key Performance Indicators Development and Their Implementation in the Production Line. *Sustainability*, *16*(7). https://doi.org/10.3390/su16072974
3. Brown, M. (2022). MES and Industry 4.0: A Synergistic Approach. *Journal of Industrial Information Integration*, *10*, 45-56.
4. Hammer, M., Champy, J. (2009). *Reengineering the Corporation: Manifesto for Business Revolution, A*. HarperCollins. https://books.google.com.bo/books?id=mjvGTXgFl6cC
5. IEEE Standard for eXtensible Event Stream (XES) for Achieving Interoperability in Event Logs and Event Streams (2023). *IEEE Std 1849-2023 (Revision of IEEE Std 1849-2016)*, 1-55. https://doi.org/10.1109/IEEESTD.2023.10267858
6. Johnson, E. (2021). Integrating MES with ERP Systems. *International Journal of Production Research*, *59*, 567-580.
7. Rebuge, Á., Ferreira, D. (2012). Business process analysis in healthcare environments: A methodology based on process mining. *Information Systems*, *37*, 99-116.
8. Smith, J. (2020). The Impact of MES on Modern Manufacturing. *Journal of Manufacturing Systems*, *45*, 123-134.
9. Thiede, S. et al. (2018). Process Mining in Manufacturing: A Review. *Journal of Manufacturing Systems*, *49*, 1-15.
10. van der Aalst, W.M.P. (2016). *Process Mining: Data Science in Action*. Berlin/Heidelberg: Springer, https://books.google.pl/books?id=hUEGDAAAQBAJ

11. van der Aalst, W., Weijters, T., Maruster, L. (2004). Workflow mining: discovering process models from event logs. *IEEE Transactions on Knowledge and Data Engineering*, *16*(9), 1128-1142. https://doi.org/10.1109/TKDE.2004.47
12. van der Aalst, W. (2024). AI and Process Mining: Enhancing Process Analysis. *Journal of Process Mining*, *12*, 34-45.